

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 10:01:39 ; Search time 1014.74 Seconds
(without alignment)
932.058 Million cell updates/sec

Title: US-09-578-453-2
Perfect score: 20
Sequence: 1 GGACATCCCGGCGCATGTCC 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 500 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	20	100.0	20 6 AR093067	AR093067 Sequence
2	20	100.0	20 6 AR093067	AR093067 Sequence
3	20	100.0	20 6 E32518	E32518 Scavenger r
4	20	100.0	20 6 E32518	E32518 Scavenger r
5	20	100.0	20 6 AR359569	AR359569 Sequence
6	20	100.0	20 6 AR359569	AR359569 Sequence
7	20	100.0	20 6 AX339212	AX339212 Sequence
8	20	100.0	20 6 AX339212	AX339212 Sequence
9	20	100.0	20 9 HSDMP53	HSDMP53
10	20	100.0	20 9 HSDMP53	HSDMP53
11	20	100.0	24 6 AX472546	AX472546 Sequence
12	20	100.0	24 6 AX472546	AX472546 Sequence
13	20	100.0	24 6 AX476842	AX476842 Sequence
14	20	100.0	24 6 AX476842	AX476842 Sequence
15	20	100.0	24 6 AX476861	AX476861 Sequence
16	20	100.0	24 6 AX476861	AX476861 Sequence
17	20	100.0	56 6 BD272327	BD272327 Selective
18	20	100.0	56 6 BD272327	BD272327 Selective
19	20	100.0	56 6 BD272328	BD272328 Selective

20	20	100.0	56 6 BD272328	BD272328 Selective
21	18.4	92.0	637 11 BV056035	BV056035 S212P6623
22	18.4	92.0	637 11 BV056035	BV056035 S212P6623
23	18.4	92.0	146698 2 AC021460	AC021460 Homo sapi
24	18.4	92.0	146698 2 AC021460	AC021460 Homo sapi
25	18.4	92.0	159478 2 AL359638	AL359638 Homo sapi
26	18.4	92.0	159478 2 AL359638	AL359638 Homo sapi
27	18.4	92.0	165090 9 HS44920	HS44920 Homo sapi
28	18.4	92.0	165090 9 HS44920	HS44920 Homo sapi
29	18.4	92.0	170132 9 AC126323	AC126323 Homo sapi
30	18.4	92.0	170132 9 AC126323	AC126323 Homo sapi
31	18.4	92.0	173987 10 AL670675	AL670675 Mouse DNA
32	18.4	92.0	173987 10 AL670675	AL670675 Mouse DNA
33	18.4	92.0	185592 10 AL773522	AL773522 Mouse DNA
34	18.4	92.0	185592 10 AL773522	AL773522 Mouse DNA
35	18.4	92.0	191128 2 AC139327	AC139327 Mus muscu
36	18.4	92.0	191128 2 AC139327	AC139327 Mus muscu
37	18.4	92.0	202227 2 AC068126	AC068126 Homo sapi
38	18.4	92.0	202227 2 AC068126	AC068126 Homo sapi
39	18.4	92.0	203950 2 AC087483	AC087483 Homo sapi
40	18.4	92.0	203950 2 AC087483	AC087483 Homo sapi
41	18.4	92.0	218688 10 AC132452	AC132452 Mus muscu
42	18.4	92.0	218688 10 AC132452	AC132452 Mus muscu
43	18.4	92.0	260120 10 AC113311	AC113311 Mus muscu
44	18.4	92.0	260120 10 AC113311	AC113311 Mus muscu
45	18.4	90.0	233077 10 AL807771	AL807771 Mouse DNA
46	18.4	90.0	233077 10 AL807771	AL807771 Mouse DNA
47	18.4	90.0	263507 2 AC093466	AC093466 Mus muscu
48	18.4	90.0	263507 2 AC093466	AC093466 Mus muscu
49	17.4	87.0	1439 1 AY395415	AY395415 Unculture
50	17.4	87.0	1439 1 AY395415	AY395415 Unculture
51	17.4	87.0	144301 9 AC010467	AC010467 Homo sapi
52	17.4	87.0	144301 9 AC010467	AC010467 Homo sapi
53	17.4	87.0	165662 9 AC092070	AC092070 Homo sapi
54	17.4	87.0	165662 9 AC092070	AC092070 Homo sapi
55	17.4	85.0	551 11 G88857	G88857 S208P644FB
56	17.4	85.0	551 11 G88857	G88857 S208P644FB
57	16.8	84.0	30 6 AR064395	AR064395 Sequence
58	16.8	84.0	30 6 AR064395	AR064395 Sequence
59	16.8	84.0	30 6 AR123186	AR123186 Sequence
60	16.8	84.0	30 6 AR123186	AR123186 Sequence
61	16.8	84.0	30 6 AR157705	AR157705 Sequence
62	16.8	84.0	30 6 AR157705	AR157705 Sequence
63	16.8	84.0	30 6 AR157705	AR157705 Sequence
64	16.8	84.0	30 6 AR157705	AR157705 Sequence
65	16.8	84.0	30 6 AR157705	AR157705 Sequence
66	16.8	84.0	30 6 AR157705	AR157705 Sequence
67	16.8	84.0	30 6 AR157705	AR157705 Sequence
68	16.8	84.0	30 6 AR157705	AR157705 Sequence
69	16.8	84.0	30 6 AR157705	AR157705 Sequence
70	16.8	84.0	30 6 AR157705	AR157705 Sequence
71	16.8	84.0	30 6 AR157705	AR157705 Sequence
72	16.8	84.0	30 6 AR157705	AR157705 Sequence
73	16.8	84.0	30 6 AR157705	AR157705 Sequence
74	16.8	84.0	30 6 AR157705	AR157705 Sequence
75	16.8	84.0	30 6 AR157705	AR157705 Sequence
76	16.8	84.0	30 6 AR157705	AR157705 Sequence
77	16.8	84.0	30 6 AR157705	AR157705 Sequence
78	16.8	84.0	30 6 AR157705	AR157705 Sequence
79	16.8	84.0	30 6 AR157705	AR157705 Sequence
80	16.8	84.0	30 6 AR157705	AR157705 Sequence
81	16.8	84.0	30 6 AR157705	AR157705 Sequence
82	16.8	84.0	30 6 AR157705	AR157705 Sequence
83	16.8	84.0	30 6 AR157705	AR157705 Sequence
84	16.8	84.0	30 6 AR157705	AR157705 Sequence
85	16.8	84.0	30 6 AR157705	AR157705 Sequence
86	16.8	84.0	30 6 AR157705	AR157705 Sequence
87	16.8	84.0	30 6 AR157705	AR157705 Sequence
88	16.8	84.0	30 6 AR157705	AR157705 Sequence
89	16.8	84.0	30 6 AR157705	AR157705 Sequence
90	16.8	84.0	30 6 AR157705	AR157705 Sequence
91	16.8	84.0	30 6 AR157705	AR157705 Sequence
92	16.8	84.0	30 6 AR157705	AR157705 Sequence

93	16.8	84.0	3328	10	BC005512	BC005512 Mus muscu	c 166	16.8	84.0	79925	9	AC093719	AC093719 Homo sapi
94	16.8	84.0	3328	10	BC005512	BC005512 Mus muscu	c 167	16.8	84.0	82561	2	AC139202	AC139202 Mus muscu
95	16.8	84.0	9391	6	AR409087	AR409087 Sequence	c 168	16.8	84.0	82561	2	AC139202	AC139202 Mus muscu
96	16.8	84.0	9391	6	AR409087	AR409087 Sequence	c 169	16.8	84.0	82646	10	AL928620	AL928620 Mouse DNA
97	16.8	84.0	9391	6	AX045103	AX045103 Sequence	c 170	16.8	84.0	82646	10	AL928620	AL928620 Mouse DNA
98	16.8	84.0	9391	6	AX045103	AX045103 Sequence	c 171	16.8	84.0	83137	10	AL844573	AL844573 Mouse DNA
99	16.8	84.0	9511	6	AR409086	AR409086 Sequence	c 172	16.8	84.0	83137	10	AL844573	AL844573 Mouse DNA
100	16.8	84.0	9511	6	AR409086	AR409086 Sequence	c 173	16.8	84.0	88526	10	AL807759	AL807759 Mouse DNA
101	16.8	84.0	9511	6	AX045101	AX045101 Sequence	c 174	16.8	84.0	88526	10	AL807759	AL807759 Mouse DNA
102	16.8	84.0	9511	6	AX045101	AX045101 Sequence	c 175	16.8	84.0	90129	10	AL627078	AL627078 Mouse DNA
103	16.8	84.0	9511	10	MMU12147	MMU12147 Mus musculu	c 176	16.8	84.0	90129	10	AL627078	AL627078 Mouse DNA
104	16.8	84.0	9511	10	MMU12147	MMU12147 Mus musculu	c 177	16.8	84.0	93553	10	AL646047	AL646047 Mouse DNA
105	16.8	84.0	15167	2	AC007412	AC007412 Mus muscu	c 178	16.8	84.0	93553	10	AL646047	AL646047 Mouse DNA
106	16.8	84.0	15167	2	AC007412	AC007412 Mus muscu	c 179	16.8	84.0	96727	10	AL627426	AL627426 Mouse DNA
107	16.8	84.0	21565	6	AX695446	AX695446 Sequence	c 180	16.8	84.0	96727	10	AL627426	AL627426 Mouse DNA
108	16.8	84.0	21565	6	AX695446	AX695446 Sequence	c 181	16.8	84.0	96916	10	AL670100	AL670100 Mouse DNA
109	16.8	84.0	22700	2	AC020792	AC020792 Mus muscu	c 182	16.8	84.0	96916	10	AL670100	AL670100 Mouse DNA
110	16.8	84.0	22700	2	AC020792	AC020792 Mus muscu	c 183	16.8	84.0	97325	10	AL928836	AL928836 Mouse DNA
111	16.8	84.0	27688	2	AC146979	AC146979 Mus muscu	c 184	16.8	84.0	97325	10	AL928836	AL928836 Mouse DNA
112	16.8	84.0	27688	2	AC146979	AC146979 Mus muscu	c 185	16.8	84.0	99268	9	AC091178	AC091178 Homo sapi
113	16.8	84.0	30210	2	AC100104	AC100104 Mus muscu	c 186	16.8	84.0	99268	9	AC091178	AC091178 Homo sapi
114	16.8	84.0	30210	2	AC100104	AC100104 Mus muscu	c 187	16.8	84.0	99606	2	AP001871	AP001871 Homo sapi
115	16.8	84.0	31638	1	RNEXFGNS	RNEXFGNS R.melliofi	c 188	16.8	84.0	99606	2	AP001871	AP001871 Homo sapi
116	16.8	84.0	31638	1	RNEXFGNS	RNEXFGNS R.melliofi	c 189	16.8	84.0	100000	9	AP000508	AP000508 Homo sapi
117	16.8	84.0	40693	10	BX784400	BX784400 Mouse DNA	c 190	16.8	84.0	100000	9	AP000508	AP000508 Homo sapi
118	16.8	84.0	40693	10	BX784400	BX784400 Mouse DNA	c 191	16.8	84.0	100803	10	AF367966	AF367966 Mus muscu
119	16.8	84.0	41506	9	AC004205	AC004205 Homo sapi	c 192	16.8	84.0	100803	10	AF367966	AF367966 Mus muscu
120	16.8	84.0	41506	9	AC004205	AC004205 Homo sapi	c 193	16.8	84.0	101563	10	AL807817	AL807817 Mouse DNA
121	16.8	84.0	42477	2	AC068805	AC068805 Mus muscu	c 194	16.8	84.0	101563	10	AL807817	AL807817 Mouse DNA
122	16.8	84.0	42477	2	AC068805	AC068805 Mus muscu	c 195	16.8	84.0	103842	10	AL840633	AL840633 Mouse DNA
123	16.8	84.0	42601	9	AC004210	AC004210 Homo sapi	c 196	16.8	84.0	103842	10	AL840633	AL840633 Mouse DNA
124	16.8	84.0	42601	9	AC004210	AC004210 Homo sapi	c 197	16.8	84.0	104509	10	AL928577	AL928577 Mouse DNA
125	16.8	84.0	43003	9	AL845556	AL845556 Human DNA	c 198	16.8	84.0	104509	10	AL928577	AL928577 Mouse DNA
126	16.8	84.0	43003	9	AL845556	AL845556 Human DNA	c 199	16.8	84.0	104746	2	AC022698	AC022698 Mus muscu
127	16.8	84.0	47355	10	AL672214	AL672214 Mouse DNA	c 200	16.8	84.0	104746	2	AC022698	AC022698 Mus muscu
128	16.8	84.0	47355	10	AL672214	AL672214 Mouse DNA	c 201	16.8	84.0	104892	10	AE008683_3	AE008683_3 Continnuation (4 of
129	16.8	84.0	49445	2	AC099873	AC099873 Mus muscu	c 202	16.8	84.0	104892	10	AE008683_3	AE008683_3 Continnuation (4 of
130	16.8	84.0	49445	2	AC099873	AC099873 Mus muscu	c 203	16.8	84.0	106635	2	AY555279	AY555279 Mus muscu
131	16.8	84.0	56934	2	AP000582	AP000582 Homo sapi	c 204	16.8	84.0	106635	2	AY555279	AY555279 Mus muscu
132	16.8	84.0	56934	2	AP000582	AP000582 Homo sapi	c 205	16.8	84.0	107898	10	AL627349	AL627349 Mouse DNA
133	16.8	84.0	59364	2	AC102347	AC102347 Mus muscu	c 206	16.8	84.0	107898	10	AL627349	AL627349 Mouse DNA
134	16.8	84.0	59364	2	AC102347	AC102347 Mus muscu	c 207	16.8	84.0	110000	1	RME591985_09	RME591985_09 Continnuation (10 o
135	16.8	84.0	60327	6	AX695572	AX695572 Sequence	c 208	16.8	84.0	110000	2	BM294366_2	BM294366_2 Continnuation (3 of
136	16.8	84.0	60327	6	AX695572	AX695572 Sequence	c 209	16.8	84.0	110000	2	BM294366_2	BM294366_2 Continnuation (3 of
137	16.8	84.0	61157	2	AC137956	AC137956 Mus muscu	c 210	16.8	84.0	110000	2	AC101867_2	AC101867_2 Continnuation (3 of
138	16.8	84.0	61157	2	AC137956	AC137956 Mus muscu	c 211	16.8	84.0	110000	2	AC101867_2	AC101867_2 Continnuation (3 of
139	16.8	84.0	62897	2	AC101230	AC101230 Mus muscu	c 212	16.8	84.0	110000	2	AL954691_1	AL954691_1 Continnuation (2 of
140	16.8	84.0	62897	2	AC101230	AC101230 Mus muscu	c 213	16.8	84.0	110000	2	AL954691_1	AL954691_1 Continnuation (2 of
141	16.8	84.0	62988	2	AC113449	AC113449 Mus muscu	c 214	16.8	84.0	110000	2	BM294366_2	BM294366_2 Continnuation (3 of
142	16.8	84.0	62988	2	AC113449	AC113449 Mus muscu	c 215	16.8	84.0	110000	10	AY363102S2_1	AY363102S2_1 Continnuation (2 of
143	16.8	84.0	63543	2	AC100544	AC100544 Mus muscu	c 216	16.8	84.0	110000	10	AY363102S2_1	AY363102S2_1 Continnuation (2 of
144	16.8	84.0	63543	2	AC100544	AC100544 Mus muscu	c 217	16.8	84.0	110000	10	AE008683_2	AE008683_2 Continnuation (3 of
145	16.8	84.0	64700	2	AC102043	AC102043 Mus muscu	c 218	16.8	84.0	110000	10	AE008683_2	AE008683_2 Continnuation (3 of
146	16.8	84.0	64700	2	AC102043	AC102043 Mus muscu	c 219	16.8	84.0	110892	9	AP000633	AP000633 Homo sapi
147	16.8	84.0	66116	2	AC116518	AC116518 Mus muscu	c 220	16.8	84.0	110892	9	AP000633	AP000633 Homo sapi
148	16.8	84.0	66116	2	AC116518	AC116518 Mus muscu	c 221	16.8	84.0	113581	2	AC109191	AC109191 Mus muscu
149	16.8	84.0	66929	2	AC121287	AC121287 Mus muscu	c 222	16.8	84.0	113581	2	AC109191	AC109191 Mus muscu
150	16.8	84.0	66929	2	AC121287	AC121287 Mus muscu	c 223	16.8	84.0	116320	9	BM248310	BM248310 Human DNA
151	16.8	84.0	68233	6	AX695956	AX695956 Sequence	c 224	16.8	84.0	116320	9	BM248310	BM248310 Human DNA
152	16.8	84.0	68233	6	AX695956	AX695956 Sequence	c 225	16.8	84.0	117000	10	AC087871	AC087871 Genomic s
153	16.8	84.0	69440	2	AC118034	AC118034 Mus muscu	c 226	16.8	84.0	117000	10	AC087871	AC087871 Genomic s
154	16.8	84.0	69440	2	AC118034	AC118034 Mus muscu	c 227	16.8	84.0	120197	9	AL671883	AL671883 Human DNA
155	16.8	84.0	70137	10	AL713988	AL713988 Mouse DNA	c 228	16.8	84.0	120197	9	AL671883	AL671883 Human DNA
156	16.8	84.0	70137	10	AL713988	AL713988 Mouse DNA	c 229	16.8	84.0	124933	10	AF545858	AF545858 Mus muscu
157	16.8	84.0	70194	2	AC133888	AC133888 Mus muscu	c 230	16.8	84.0	124933	10	AF545858	AF545858 Mus muscu
158	16.8	84.0	70194	2	AC133888	AC133888 Mus muscu	c 231	16.8	84.0	125641	10	AF545858	AF545858 Mus muscu
159	16.8	84.0	70991	10	AY225417	AY225417 Mus muscu	c 232	16.8	84.0	125641	10	AF545858	AF545858 Mus muscu
160	16.8	84.0	70991	10	AY225417	AY225417 Mus muscu	c 233	16.8	84.0	127025	2	AC125522	AC125522 Mus muscu
161	16.8	84.0	75039	2	AC102185	AC102185 Mus muscu	c 234	16.8	84.0	127025	2	AC125522	AC125522 Mus muscu
162	16.8	84.0	75039	2	AC102185	AC102185 Mus muscu	c 235	16.8	84.0	127380	10	AC125113	AC125113 Mus muscu
163	16.8	84.0	78135	2	AC139059	AC139059 Mus muscu	c 236	16.8	84.0	127380	10	AC125113	AC125113 Mus muscu
164	16.8	84.0	78135	2	AC139059	AC139059 Mus muscu	c 237	16.8	84.0	129451	2	AC073783	AC073783 Mus muscu
165	16.8	84.0	79925	9	AC093719	AC093719 Homo sapi	c 238	16.8	84.0	129451	2	AC073783	AC073783 Mus muscu

239	15.8	84.0	130301	10	EX813328	Mouse	DNA	C 312	16.8	84.0	150643	2	AP001844
C 240	16.8	84.0	130301	10	EX813328	Mouse	DNA	C 313	16.8	84.0	151639	2	AP001193
C 241	16.8	84.0	130474	2	AC079039	Mus	muscu	C 314	16.8	84.0	151639	2	AP001193
C 242	16.8	84.0	130474	2	AC079039	Mus	muscu	C 315	16.8	84.0	152275	2	AC147129
C 243	16.8	84.0	1311732	10	AP259071	Mus	muscu	C 316	16.8	84.0	152275	2	AC147129
C 244	16.8	84.0	1311732	10	AP259071	Mus	muscu	C 317	16.8	84.0	153847	10	AL591763
C 245	16.8	84.0	132342	10	AC132566	Mus	muscu	C 318	16.8	84.0	153847	10	AL591763
C 246	16.8	84.0	132342	10	AC132566	Mus	muscu	C 319	16.8	84.0	154373	10	AC121988
C 247	16.8	84.0	132668	2	AC099829	Homo	sapi	C 320	16.8	84.0	154373	10	AC121988
C 248	16.8	84.0	132668	2	AC099829	Homo	sapi	C 321	16.8	84.0	154754	2	AC105982
C 249	16.8	84.0	132688	2	AC119878	Mus	muscu	C 322	16.8	84.0	154754	2	AC105982
C 250	16.8	84.0	132688	2	AC119878	Mus	muscu	C 323	16.8	84.0	155954	10	AL298860
C 251	16.8	84.0	135432	2	AL645589	Mus	muscu	C 324	16.8	84.0	155954	10	AL298860
C 252	16.8	84.0	135432	2	AL645589	Mus	muscu	C 325	16.8	84.0	156468	10	AL954817
C 253	16.8	84.0	135453	2	AC118736	Mus	muscu	C 326	16.8	84.0	156468	10	AL954817
C 254	16.8	84.0	135453	2	AC118736	Mus	muscu	C 327	16.8	84.0	156603	10	AC120787
C 255	16.8	84.0	136155	10	AC132586	Mus	muscu	C 328	16.8	84.0	156603	10	AC120787
C 256	16.8	84.0	136155	10	AC132586	Mus	muscu	C 329	16.8	84.0	157019	10	AL772156
C 257	16.8	84.0	137909	2	AC117640	Mus	muscu	C 330	16.8	84.0	157019	10	AL772156
C 258	16.8	84.0	137909	2	AC117640	Mus	muscu	C 331	16.8	84.0	157055	10	AC134524
C 259	16.8	84.0	139160	2	AC131230	Mus	muscu	C 332	16.8	84.0	157055	10	AC134524
C 260	16.8	84.0	139160	2	AC131230	Mus	muscu	C 333	16.8	84.0	157958	10	AC1347953
C 261	16.8	84.0	140494	10	EX276179	Mouse	DNA	C 334	16.8	84.0	157958	10	AC147983
C 262	16.8	84.0	140494	10	EX276179	Mouse	DNA	C 335	16.8	84.0	158006	10	AC127687
C 263	16.8	84.0	141036	2	AC023807	Mus	muscu	C 336	16.8	84.0	158006	10	AC127687
C 264	16.8	84.0	141036	2	AC023807	Mus	muscu	C 337	16.8	84.0	158009	10	AL837508
C 265	16.8	84.0	141324	10	AL645466	Mus	DNA	C 338	16.8	84.0	158009	10	AL837508
C 266	16.8	84.0	141324	10	AL645466	Mus	DNA	C 339	16.8	84.0	158764	10	AC132083
C 267	16.8	84.0	141678	10	AC102722	Mus	muscu	C 340	16.8	84.0	158764	10	AC132083
C 268	16.8	84.0	141678	10	AC102722	Mus	muscu	C 341	16.8	84.0	159669	2	AC102017
C 269	16.8	84.0	142565	10	AL7772205	Mouse	DNA	C 342	16.8	84.0	159669	2	AC102017
C 270	16.8	84.0	142565	10	AL7772205	Mouse</							

385	16.8	84.0	164866	2	AC091102	Homo sapi
386	16.8	84.0	164866	2	AC091102	Homo sapi
387	16.8	84.0	165703	10	AC139296	Mus muscu
388	16.8	84.0	165703	10	AC139296	Mus muscu
389	16.8	84.0	165733	2	AC107706	Mus muscu
390	16.8	84.0	165733	2	AC107706	Mus muscu
391	16.8	84.0	166042	9	AP006289	Homo sapi
392	16.8	84.0	166042	9	AP006289	Homo sapi
393	16.8	84.0	166231	2	AC087037	Mus muscu
394	16.8	84.0	166231	2	AC087037	Mus muscu
395	16.8	84.0	166232	2	AC135667	Mus muscu
396	16.8	84.0	166232	2	AC135667	Mus muscu
397	16.8	84.0	166335	2	AC073799	Mus muscu
398	16.8	84.0	166335	2	AC073799	Mus muscu
399	16.8	84.0	167166	10	AL732520	Mouse DNA
400	16.8	84.0	167166	10	AL732520	Mouse DNA
401	16.8	84.0	167418	10	AL845294	Mouse DNA
402	16.8	84.0	167418	10	AL845294	Mouse DNA
403	16.8	84.0	168105	2	AC115069	Mus muscu
404	16.8	84.0	168105	2	AC115069	Mus muscu
405	16.8	84.0	168117	10	AC132352	Mus muscu
406	16.8	84.0	168117	10	AC132352	Mus muscu
407	16.8	84.0	168251	2	AC112940	Mus muscu
408	16.8	84.0	168251	2	AC112940	Mus muscu
409	16.8	84.0	168515	10	AC130215	Mus muscu
410	16.8	84.0	168515	10	AC130215	Mus muscu
411	16.8	84.0	168698	2	AC127583	Mus muscu
412	16.8	84.0	168698	2	AC127583	Mus muscu
413	16.8	84.0	168916	2	AC124822	Mus muscu
414	16.8	84.0	168916	2	AC124822	Mus muscu
415	16.8	84.0	169247	10	AC103621	Mus muscu
416	16.8	84.0	169247	10	AC103621	Mus muscu
417	16.8	84.0	169481	10	AC122403	Mus muscu
418	16.8	84.0	169481	10	AC122403	Mus muscu
419	16.8	84.0	169815	10	AC122057	Mus muscu
420	16.8	84.0	169815	10	AC122057	Mus muscu
421	16.8	84.0	170020	9	EX072579	Human DNA
422	16.8	84.0	170020	9	EX072579	Human DNA
423	16.8	84.0	170119	2	AC143515	Macaca mu
424	16.8	84.0	170119	2	AC143515	Macaca mu
425	16.8	84.0	170369	2	AC102188	Mus muscu
426	16.8	84.0	170369	2	AC102188	Mus muscu
427	16.8	84.0	170548	10	AC074041	Mus muscu
428	16.8	84.0	170548	10	AC074041	Mus muscu
429	16.8	84.0	171122	10	AC124394	Mus muscu
430	16.8	84.0	171122	10	AC124394	Mus muscu
431	16.8	84.0	171209	10	AC132622	Mus muscu
432	16.8	84.0	171209	10	AC132622	Mus muscu
433	16.8	84.0	171261	10	AC110514	Mus muscu
434	16.8	84.0	171261	10	AC110514	Mus muscu
435	16.8	84.0	171527	2	AC129942	Mus muscu
436	16.8	84.0	171527	2	AC129942	Mus muscu
437	16.8	84.0	172237	10	AC134475	Mus muscu
438	16.8	84.0	172237	10	AC134475	Mus muscu
439	16.8	84.0	172263	2	AC119280	Mus muscu
440	16.8	84.0	172263	2	AC119280	Mus muscu
441	16.8	84.0	172692	5	EX005380	Zebrafish
442	16.8	84.0	172692	5	EX005380	Zebrafish
443	16.8	84.0	172694	2	AC124116	Mus muscu
444	16.8	84.0	172694	2	AC124116	Mus muscu
445	16.8	84.0	173181	10	AC144525	Mus muscu
446	16.8	84.0	173181	10	AC144525	Mus muscu
447	16.8	84.0	173351	10	AL732502	Mouse DNA
448	16.8	84.0	173351	10	AL732502	Mouse DNA
449	16.8	84.0	173769	10	AL669819	Mouse DNA
450	16.8	84.0	173769	10	AL669819	Mouse DNA
451	16.8	84.0	173926	2	MM454K24	Mus muscu
452	16.8	84.0	173926	2	MM454K24	Mus muscu
453	16.8	84.0	174670	2	AC112933	Mus muscu
454	16.8	84.0	174670	2	AC112933	Mus muscu
455	16.8	84.0	174750	10	AC124359	Mus muscu
456	16.8	84.0	174750	10	AC124359	Mus muscu
457	16.8	84.0	174928	10	AC121975	Mus muscu

458	16.8	84.0	174928	10	AC121975	Mus muscu
459	16.8	84.0	175092	10	AC122491	Mus muscu
460	16.8	84.0	175092	10	AC122491	Mus muscu
461	16.8	84.0	175131	10	AC121821	Mus muscu
462	16.8	84.0	175131	10	AC121821	Mus muscu
463	16.8	84.0	175316	2	AC136516	Mus muscu
464	16.8	84.0	175316	2	AC136516	Mus muscu
465	16.8	84.0	175340	2	AC110159	Mus muscu
466	16.8	84.0	175340	2	AC110159	Mus muscu
467	16.8	84.0	175809	10	AL714027	Mouse DNA
468	16.8	84.0	175809	10	AL714027	Mouse DNA
469	16.8	84.0	175878	2	AC122554	Mus muscu
470	16.8	84.0	175878	2	AC122554	Mus muscu
471	16.8	84.0	175987	10	AC079273	Mus muscu
472	16.8	84.0	175987	10	AC079273	Mus muscu
473	16.8	84.0	176056	10	AC073437	Mus muscu
474	16.8	84.0	176056	10	AC073437	Mus muscu
475	16.8	84.0	176278	10	AC126557	Mus muscu
476	16.8	84.0	176278	10	AC126557	Mus muscu
477	16.8	84.0	176284	9	AB054536	Pan trogl
478	16.8	84.0	176284	9	AB054536	Pan trogl
479	16.8	84.0	176304	2	AC102182	Mus muscu
480	16.8	84.0	176304	2	AC102182	Mus muscu
481	16.8	84.0	176436	10	AC129937	Mus muscu
482	16.8	84.0	176436	10	AC129937	Mus muscu
483	16.8	84.0	176541	10	AC127367	Mus muscu
484	16.8	84.0	176541	10	AC127367	Mus muscu
485	16.8	84.0	176711	10	AC137969	Mus muscu
486	16.8	84.0	176711	10	AC137969	Mus muscu
487	16.8	84.0	176871	10	AC124687	Mus muscu
488	16.8	84.0	176871	10	AC124687	Mus muscu
489	16.8	84.0	177183	2	AC112973	Mus muscu
490	16.8	84.0	177183	2	AC112973	Mus muscu
491	16.8	84.0	177256	10	AL691476	Mouse DNA
492	16.8	84.0	177256	10	AL691476	Mouse DNA
493	16.8	84.0	177302	10	AC127415	Mus muscu
494	16.8	84.0	177302	10	AC127415	Mus muscu
495	16.8	84.0	177450	2	AC103674	Mus muscu
496	16.8	84.0	177450	2	AC103674	Mus muscu
497	16.8	84.0	177788	2	AC069310	Mus muscu
498	16.8	84.0	177788	2	AC069310	Mus muscu
499	16.8	84.0	177955	2	AC131034	Mus muscu
500	16.8	84.0	177955	2	AC131034	Mus muscu

ALIGNMENTS

RESULT 1	AR093067	20 bp	DNA	linear	PAT 08-SEP-2000
LOCUS	Sequence 162 from patent US 5998383.				
DEFINITION	AR093067				
ACCESSION	AR093067.1	GI:10019819			
VERSION					
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 20)				
AUTHORS	Wright,J.A. and Young,A.H.				
TITLE	Antitumor antiense sequences directed against ribonucleotide				
JOURNAL	Patent: US 5998383-A 162 07-DEC-1999;				
FEATURES	Location/Qualifiers				
source	1..20				
ORIGIN	/organism="unknown"				
	/mol_type="unassigned DNA"				

Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Fred. No. 73;
 Matches 20; Conservative 0; Mismatches 0; Gaps 0;

Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCCGGCATGTCC 20
 DB 1 GGACATGCCCCGGCATGTCC 20

RESULT 6
 AR359569/c 20 bp DNA linear PAT 17-AUG-2003
 LOCUS AR359569
 DEFINITION Sequence 162 from patent US 6593305.
 ACCESSION AR359569
 VERSION AR359569.1 GI:33766292
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclashed.
 AUTHORS 1 (bases 1 to 20)
 TITLE Wright, J. A.
 JOURNAL Antitumor antisense sequences directed against R1 and R2 components
 FEATURES of ribonucleotide reductase
 source Patent: US 6593305-A 162 15-JUL-2003;
 Location/Qualifiers
 1..20
 /organism="Unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCCGGCATGTCC 20
 DB 20 GGACATGCCCCGGCATGTCC 1

RESULT 7
 AX339212 20 bp DNA linear PAT 10-JAN-2002
 LOCUS AX339212
 DEFINITION Sequence 6 from Patent WO0196602.
 ACCESSION AX339212
 VERSION AX339212.1 GI:18135473
 KEYWORDS
 SOURCE Synthetic construct
 ORGANISM synthetic construct
 REFERENCE artificial sequences.
 1
 AUTHORS Yang, A.L. and Festing, M.
 TITLE Methods and materials to determine the p53 status of a sample by
 JOURNAL determining the binding of p53 to a vector
 MEDLINE Patent: WO 0196602-A 6 20-DEC-2001;
 JOURNAL MEDICAL RESEARCH COUNCIL (GB)
 FEATURES Location/Qualifiers
 1..20
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Consensus sequence"

ORIGIN
 Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCCGGCATGTCC 20
 DB 1 GGACATGCCCCGGCATGTCC 20

RESULT 8

AX339212/c 20 bp DNA linear PAT 10-JAN-2002
 LOCUS AX339212
 DEFINITION Sequence 6 from Patent WO0196602.
 ACCESSION AX339212
 VERSION AX339212.1 GI:18135473
 KEYWORDS
 SOURCE Synthetic construct
 ORGANISM synthetic construct
 REFERENCE artificial sequences.
 1
 AUTHORS Yang, A.L. and Festing, M.
 TITLE Methods and materials to determine the p53 status of a sample by
 JOURNAL determining the binding of p53 to a vector
 MEDLINE Patent: WO 0196602-A 6 20-DEC-2001;
 JOURNAL MEDICAL RESEARCH COUNCIL (GB)
 FEATURES Location/Qualifiers
 1..20
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Consensus sequence"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCCGGCATGTCC 20
 DB 20 GGACATGCCCCGGCATGTCC 1

RESULT 9
 HSDMP53 20 bp DNA linear PRI 05-JUN-1992
 LOCUS HSDMP53
 DEFINITION H. sapiens DNA binding motif recognized by p53 protein containing
 X63571
 X63571.1 GI:30483
 ACCESSION DNA motif; high affinity DNA binding motif.
 VERSION
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Funk, W.D., Pak, D.T., Karas, R.H., Wright, W.E. and Shay, J.W.
 TITLE A transcriptionally active DNA-binding site for human p53 protein
 JOURNAL complexes
 MEDLINE Mol. Cell. Biol. 12 (6), 2866-2871 (1992)
 PUBMED 92269860
 JOURNAL 1588974
 REFERENCE 2 (bases 1 to 20)
 AUTHORS Shay, J.W.
 TITLE Direct Submission
 JOURNAL Submitted (09-JAN-1992) J.W. Shay, The University of Texas,
 Southwestern Medical Center at Dallas, Dept of Cell Biology &
 Neuroscience, 5323 Harry Hines Boulevard, Dallas TX 75235-9039,
 USA

COMMENT
 See also X63905-7 & S.E. Kern et al, Science 252:1708-1711(1991).
 This sequence is a high-affinity DNA binding motif recognized by
 p53 protein containing complexes and is likely to be a regulatory
 signal. Variations of this sequence is found in the 5' flanking
 region, 3' flanking or untranslated regions of many genes associated
 with cellular proliferation including a perfect match in the 5'
 flanking regions of the c-Ha-ras-1 oncogene, the gene for DNA
 ligase 1 & in the 3' untranslated region of the tyrosine kinase
 receptor gene. When placed upstream of a reporter construct, this
 consensus sequence promotes p53 dependent transcription.

FEATURES
 source
 1..20
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN /cell_type="diploid fibroblasts"

Query Match 100.0%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
|||||
1 GGACATGCCCGGCATGTCC 20

Db 1 GGACATGCCCGGCATGTCC 20

RESULT 10
HSDMP53/c
LOCUS
DEFINITION H sapiens DNA binding motif recognized by p53 protein containing
complexes.
X63571
X63571.1 GI:30483
VERSION DNA motif; high affinity DNA binding motif.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 20)
AUTHORS Funk, W.D., Pak, D.T., Karas, R.H., Wright, W.B. and Shay, J.W.
TITLE A transcriptionally active DNA-binding site for human p53 protein
complexes
JOURNAL Mol. Cell. Biol. 12 (6), 2866-2871 (1992)
MEDLINE 92269860
PUBMED 1588974
REFERENCE 2 (bases 1 to 20)
AUTHORS Shay, J.W.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1992) J.W. Shay, The University of Texas,
Southwestern Medical Center at Dallas, Dept of Cell Biology &
Neurosciences, 5323 Harry Hines Boulevard, Dallas TX 75235-9039,
USA

COMMENT See also X63905-7 & S.B. Kern et al. Science 252:1708-1711(1991).
This sequence is a high-affinity DNA binding motif recognized by
p53 protein containing complexes and is likely to be a regulatory
signal. Variations of this sequence is found in the 5' flanking
region, 3' flanking or untranslated regions of many genes associated
with cellular proliferation including a perfect match in the 5'
flanking regions of the c-Ha-ras-1 oncogene, the gene for DNA
ligase I & in the 3' untranslated region of a reporter construct, this
receptor gene. When placed upstream of a reporter construct, this
consensus sequence promotes p53 dependent transcription.

FEATURES
source
1..20
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/cell_type="diploid fibroblasts"

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
|||||
1 GGACATGCCCGGCATGTCC 20

Db 20 GGACATGCCCGGCATGTCC 1

RESULT 11
AX472546 24 bp DNA linear PAT 09-AUG-2002
LOCUS
DEFINITION Sequence 41 from Patent WO02052039.
ACCESSION AX472546
VERSION AX472546.1 GI:22207450
KEYWORDS
SOURCE
ORGANISM
REFERENCE
synthetic construct

ORGANISM
synthetic construct
artificial sequences.

REFERENCE
1
AUTHORS Blais, Y., Rousseau, P., Leblanc, B. and Camero, R.N.
TITLE Methods for selecting and producing selective pharmaceutical
compounds and compositions using an established genetically altered
cell-based library responsive to transcription factors; genetic
constructs and library therefor
JOURNAL Patent: WO 02052039-A 41 04-JUL-2002;
Geneka Biotechnology Inc. (CA)
FEATURES
source
1..24
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Oligonucleotide"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
|||||
5 GGACATGCCCGGCATGTCC 24

Db 5 GGACATGCCCGGCATGTCC 24

RESULT 12
AX472546/c
LOCUS AX472546 24 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 41 from Patent WO02052039.
ACCESSION AX472546
VERSION AX472546.1 GI:22207450
KEYWORDS
SOURCE
ORGANISM
REFERENCE
synthetic construct
synthetic construct
artificial sequences.

1
AUTHORS Blais, Y., Rousseau, P., Leblanc, B. and Camero, R.N.
TITLE Methods for selecting and producing selective pharmaceutical
compounds and compositions using an established genetically altered
cell-based library responsive to transcription factors; genetic
constructs and library therefor
JOURNAL Patent: WO 02052039-A 41 04-JUL-2002;
Geneka Biotechnology Inc. (CA)
FEATURES
source
1..24
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Oligonucleotide"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
|||||
24 GGACATGCCCGGCATGTCC 5

Db 24 GGACATGCCCGGCATGTCC 5

RESULT 13
AX476842 24 bp DNA linear PAT 12-AUG-2002
LOCUS
DEFINITION Sequence 19 from Patent WO02052037.
ACCESSION AX476842
VERSION AX476842.1 GI:22216118
KEYWORDS
SOURCE
ORGANISM
REFERENCE
synthetic construct
synthetic construct
artificial sequences.

AUTHORS Larose, A.M., Rousseau, P., Leb lanc, B. and Camato, R.
 TITLE Method for screening and/or identifying factors that bind to
 nucleic acids
 JOURNAL Patent: WO 02052037-A 19 04-JUL-2002;
 Geneka Biotechnology Inc. (CA)

FEATURES
 source
 1. .24
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="NABE-probes"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
 |||||
 5 GGACATGCCCGGCATGTCC 24

RESULT 14
 AX476842/c 24 bp DNA linear PAT 12-AUG-2002
 LOCUS AX476842
 DEFINITION Sequence 19 from Patent WO02052037.
 ACCESSION AX476842
 VERSION AX476842.1 GI:22216118
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 artificial sequences.

REFERENCE
 1
 AUTHORS Larose, A.M., Rousseau, P., Leb lanc, B. and Camato, R.
 TITLE Method for screening and/or identifying factors that bind to
 nucleic acids
 JOURNAL Patent: WO 02052037-A 19 04-JUL-2002;
 Geneka Biotechnology Inc. (CA)

FEATURES
 source
 1. .24
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="NABE-probes"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
 |||||
 24 GGACATGCCCGGCATGTCC 5

RESULT 15
 AX476861 24 bp DNA linear PAT 12-AUG-2002
 LOCUS AX476861
 DEFINITION Sequence 38 from Patent WO02052037.
 ACCESSION AX476861
 VERSION AX476861.1 GI:22216137
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 artificial sequences.

REFERENCE
 1
 AUTHORS Larose, A.M., Rousseau, P., Leb lanc, B. and Camato, R.
 TITLE Method for screening and/or identifying factors that bind to
 nucleic acids
 JOURNAL Patent: WO 02052037-A 38 04-JUL-2002;
 Geneka Biotechnology Inc. (CA)

FEATURES
 source
 1. .24
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="NABE-probes"

/organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Double stranded NABE"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
 |||||
 5 GGACATGCCCGGCATGTCC 24

RESULT 16
 AX476861/c 24 bp DNA linear PAT 12-AUG-2002
 LOCUS AX476861
 DEFINITION Sequence 38 from Patent WO02052037.
 ACCESSION AX476861
 VERSION AX476861.1 GI:22216137
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 artificial sequences.

REFERENCE
 1
 AUTHORS Larose, A.M., Rousseau, P., Leb lanc, B. and Camato, R.
 TITLE Method for screening and/or identifying factors that bind to
 nucleic acids
 JOURNAL Patent: WO 02052037-A 38 04-JUL-2002;
 Geneka Biotechnology Inc. (CA)

FEATURES
 source
 1. .24
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Double stranded NABE"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
 |||||
 24 GGACATGCCCGGCATGTCC 5

RESULT 17
 BD272327 56 bp DNA linear PAT 17-JUL-2003
 LOCUS BD272327
 DEFINITION Selectively replicating viral vector.
 ACCESSION BD272327
 VERSION BD272327.1 GI:33082095
 KEYWORDS JP 2002541761-A/9.
 SOURCE synthetic construct
 ORGANISM synthetic construct
 artificial sequences.

REFERENCE
 1 (bases 1 to 56)
 AUTHORS Ramachandra, M. and Shabram, P.W.
 TITLE Selectively replicating viral vector
 JOURNAL Patent: JP 2002541761-A 9 10-DEC-2002;
 CANUT INC

COMMENT
 CS Artificial Sequence
 PN JP 2002541761-A/9
 PD 10-DEC-2002 JP 2000576027
 PF 14-OCT-1999 JP 2000576027
 PR 15-OCT-1998 US 09/172686
 PI MURALIDHARA RAMACHANDRA, PAUL W SHABRAM
 PC C12N15/09, A61K35/76, A61K48/00, A61P35/00, C12N5/10, C12N7/00, PC
 C12N15/00
 PC C12N5/00
 CC complementary 5'-phosphorylated oligonucleotide containing CC

FEATURES
Source
CC binding sites(p53CON)
FH Key Location/Qualifiers
FT source 1..56
Location/Qualifiers
1..56
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 56;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
|||||
7 GGACATGCCCGGCGCATGTCC 26

Db
BD272327/c
LOCUS BD272327 56 bp DNA linear PAT 17-JUL-2003
DEFINITION Selectively replicating viral vector.
ACCESSION BD272327.1 GI:33082095
VERSION UP 2002541761-A/9.
KEYWORDS
SOURCE
ORGANISM
synthetic construct
artificial sequences.
1 (bases 1 to 56)
Ramachandra,M. and Shabram,P.W.
REFERENCE
AUTHORS
TITLE
JOURNAL
Patent: JP 2002541761-A 9 10-DEC-2002;

COMMENT
OS Artificial Sequence
PN JP 2002541761-A/9
PD 10-DEC-2002
PF 14-OCT-1999 JP 2000576027
PR 15-OCT-1998 US 09/172686
PI MURALIDHARA RAMACHANDRA, PAUL W SHABRAM
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,C12N5/10,C12N7/00, PC
C12N15/00,
PC C12N5/00
CC complementary 5'-phosphorylated oligonucleotide containing CC
consensus p53
CC binding sites(p53CON)
FH Key Location/Qualifiers
FT source 1..56
Location/Qualifiers
/organism="Artificial Sequence".

FEATURES
Source
1..56
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 56;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
|||||
51 GGACATGCCCGGCGCATGTCC 32

Db
BD272328
LOCUS BD272328 56 bp DNA linear PAT 17-JUL-2003
DEFINITION Selectively replicating viral vector.
ACCESSION BD272328.1 GI:33082096
VERSION

KEYWORDS JP 2002541761-A/10.
SOURCE
ORGANISM
synthetic construct
artificial sequences.
1 (bases 1 to 56)
Ramachandra,M. and Shabram,P.W.
REFERENCE
AUTHORS
TITLE
JOURNAL
Patent: JP 2002541761-A 10 10-DEC-2002;

COMMENT
OS Artificial Sequence
PN JP 2002541761-A/10
PD 10-DEC-2002
PF 14-OCT-1999 JP 2000576027
PR 15-OCT-1998 US 09/172686
PI MURALIDHARA RAMACHANDRA, PAUL W SHABRAM
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,C12N5/10,C12N7/00, PC
C12N15/00,
PC C12N5/00
CC complementary 5'-phosphorylated oligonucleotide containing CC
consensus p53
CC binding sites(p53CON)
FH Key Location/Qualifiers
FT source 1..56
Location/Qualifiers
/organism="Artificial Sequence".

FEATURES
Source
1..56
/organism="synthetic construct"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 56;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
|||||
2 GGACATGCCCGGCGCATGTCC 21

Db
BD272328
LOCUS BD272328 56 bp DNA linear PAT 17-JUL-2003
DEFINITION Selectively replicating viral vector.
ACCESSION BD272328.1 GI:33082096
VERSION UP 2002541761-A/10.
KEYWORDS
SOURCE
ORGANISM
synthetic construct
artificial sequences.
1 (bases 1 to 56)
Ramachandra,M. and Shabram,P.W.
REFERENCE
AUTHORS
TITLE
JOURNAL
Patent: JP 2002541761-A 10 10-DEC-2002;

COMMENT
OS Artificial Sequence
PN JP 2002541761-A/10
PD 10-DEC-2002
PF 14-OCT-1999 JP 2000576027
PR 15-OCT-1998 US 09/172686
PI MURALIDHARA RAMACHANDRA, PAUL W SHABRAM
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,C12N5/10,C12N7/00, PC
C12N15/00,
PC C12N5/00
CC complementary 5'-phosphorylated oligonucleotide containing CC
consensus p53
CC binding sites(p53CON)
FH Key Location/Qualifiers
FT source 1..56
Location/Qualifiers
/organism="Artificial Sequence".

FEATURES
Source
1..56
/organism="synthetic construct"

ORIGIN /mol_type="genomic DNA"
/db_xref="taxon:32630"

Query Match 100.0%; Score 20; DB 6; Length 56;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 46 GGACATGCCCGGCGCATGTCC 27

RESULT 21
LOCUS BV056035
DEFINITION S212P6623PB2.T0 CZECHII/E1 Mus musculus STS genomic, sequence
ACCESSION BV056035
VERSION BV056035.1 GI:31171830
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
TITLE The mosaic structure of variation in the laboratory mouse genome
JOURNAL Nature 420 (6915), 574-578 (2002)
MEDLINE 22354684
COMMENT 12466852

Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 637

Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SVIMJ, C3H/HeJ, and BALB/cBYJ. The WGS
reads were placed uniquely on the MGSCV3 C57BL/6J assembly and SNP
detection was carried out by SSNA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

FEATURES
source location/Qualifiers

1..637
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CZECHII/E1"
/db_xref="taxon:10090"
/map="11 23-568 3142702-3143247"
/clone_11b="CZECHII/E1"
1..5637

STS

Query Match 92.0%; Score 18.4; DB 11; Length 637;
Best Local Similarity 95.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 497 GGACATGCCCGGCGCATGTCC 516

RESULT 22

BV056035/c 637 bp DNA linear STS 31-MAY-2003
LOCUS S212P6623PB2.T0 CZECHII/E1 Mus musculus STS genomic, sequence
DEFINITION tagged site.

ACCESSION BV056035
VERSION BV056035.1 GI:31171830
KEYWORDS STS.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.

REFERENCE The mosaic structure of variation in the laboratory mouse genome
JOURNAL Nature 420 (6915), 574-578 (2002)
MEDLINE 22354684
COMMENT 12466852

Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 637

Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SVIMJ, C3H/HeJ, and BALB/cBYJ. The WGS
reads were placed uniquely on the MGSCV3 C57BL/6J assembly and SNP
detection was carried out by SSNA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

FEATURES
source location/Qualifiers

1..637
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CZECHII/E1"
/db_xref="taxon:10090"
/map="11 23-568 3142702-3143247"
/clone_11b="CZECHII/E1"
1..5637

STS

Query Match 92.0%; Score 18.4; DB 11; Length 637;
Best Local Similarity 95.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 516 GGACATGCCCGGCGCATGTCC 497

RESULT 23
LOCUS AC021460 146698 bp DNA linear PRI 15-MAR-2002
DEFINITION Homo sapiens chromosome 15, clone RP11-259F15, complete sequence.
ACCESSION AC021460
VERSION AC021460.10 GI:19482349
KEYWORDS HUG.

SOURCE Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 15, clone RP11-259F15
JOURNAL Unpublished

```

REFERENCE
AUTHORS      2 (bases 1 to 146698)
              Birtten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
              Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
              Boguslavsky,L., Bouhgalter,B., Brown,A., Burkett,G., Castle,A.,
              Choebel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
              Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fensholt,J.,
              Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
              Gardina,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
              Howland,J.C., Johnson,R., Jones,C., Kam,L., Karatas,A., Klein,J.,
              Lander,T., Lehoczy,J., Levine,R., Lien,C., Liu,G., Locke,K.,
              Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
              McSheeher,R., Meldrim,J., Menus,L., Morrow,J., Naylor,J.,
              Norman,C.H., O'Connor,T., O'Donnell,P., Olyvar,T.M., Peterson,K.,
              Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
              Roy,A., Santos,R., Severy,P., Spencer,B., Strange-Thomann,N.,
              Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
              Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
              Zimmer,A. and Zody,M.

TITLE
JOURNAL      Direct Submission
REFERENCE    Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
AUTHORS      Research, 320 Charles Street, Cambridge, MA 02141, USA
              3 (bases 1 to 146698)
              Birtten,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
              Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
              Bouhgalter,B., Brown,A., Camarata,J., Campoliano,A., Chang,J.,
              Chazaro,B., Choebel,Y., Colangelo,M., Collins,S., Collymore,A.,
              Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardina,S.,
              Glinde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
              Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
              Kanat,A., Karatas,A., Kelle,C., Lacroque,K., Lamazares,R.,
              Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
              Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
              McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Menus,L.,
              Milnova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
              Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
              Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N., Pollara,V.,
              Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
              Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnapack,R.,
              Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
              Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
              Tophan,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
              Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
              Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL      Direct Submission
COMMENT      Submitted (15-MAR-2002) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              On Mar 15, 2002 this sequence version replaced gi:17223355.
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html

              Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence.submissions@genome.wi.mit.edu

              Project Information
              Center project name: L5195
              Center clone name: 259_F_15

FEATURES
SOURCE
      1. 146698
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="15"
      /map="15"
      /clone="RP11-259P15"
      /clone.lib="RP11-11 Human Male BAC"
      59. 119
      /rpt_family="MIR"
      repeat_region
      216. 561
      /rpt_family="AluDb"
      repeat_region
  
```

```

repeat_region complement(562. .732)
/rpt_family="L2"
repeat_region 1082. .119
/rpt_family="AT_rich"
repeat_region 1120. .1415
/rpt_family="AluDb"
repeat_region complement(1417. .1696)
/rpt_family="L1MB8"
repeat_region complement(1697. .1985)
/rpt_family="AluSc"
repeat_region complement(1986. .2138)
/rpt_family="L1MB8"
repeat_region 2798. .2881
/rpt_family="MIR"
repeat_region 2958. .3108
/rpt_family="L1TR33"
repeat_region complement(3113. .3376)
/rpt_family="L1PBA"
repeat_region 3397. .3639
/rpt_family="L1TR33A"
repeat_region complement(3640. .3977)
/rpt_family="THE1B"
repeat_region 3978. .4097
/rpt_family="L1TR33A"
repeat_region complement(4257. .4566)
/rpt_family="AluSc"
repeat_region complement(4971. 5131)
/rpt_family="FRAM"
repeat_region 5255. .5494
/rpt_family="AluDo"
repeat_region 5649. .5963
/rpt_family="AluSc"
repeat_region 6659. .6791
/rpt_family="AluDo"
repeat_region 8068. .8244
/rpt_family="AluDo"
repeat_region 9323. .9497
/rpt_family="L3"
repeat_region 9600. .9871
/rpt_family="L3"
repeat_region 9875. .10038
/rpt_family="AluDo"
repeat_region 10045. .10340
/rpt_family="AluSc"
repeat_region 10371. .10419
/rpt_family="AT_rich"
repeat_region complement(10479. .10844)
/rpt_family="L2"
repeat_region complement(11081. .11323)
/rpt_family="L1MC4a"
repeat_region complement(11750. .12055)
/rpt_family="AluSc"
repeat_region 12281. .12405
/rpt_family="MIR"
repeat_region 12691. .12764
/rpt_family="TATATG)n"
repeat_region 13182. .13457
/rpt_family="AluSc"
repeat_region 13461. .13484
/rpt_family="(CAA)n"
repeat_region 13568. .13651
/rpt_family="MIR3"
repeat_region complement(13859. .14002)
/rpt_family="MIR3"
repeat_region complement(14683. .14795)
/rpt_family="MIR"
repeat_region 14892. .15185
/rpt_family="AluSc"
repeat_region complement(16265. .16367)
/rpt_family="MIR"
repeat_region 16368. .16405
/rpt_family="(TG)n"
repeat_region complement(16406. .16541)

```



```

repeat_region      /rpt_family="MIR"
                    complement(16553..16824)
repeat_region      /rpt_family="L1MC4a"
                    complement(16974..17273)
repeat_region      /rpt_family="AluSx"
                    complement(17892..18185)
repeat_region      /rpt_family="AluSx"
                    complement(19150..19457)
repeat_region      /rpt_family="AluY"
                    complement(19464..19526)
repeat_region      /rpt_family="MER5A"
                    complement(19527..19922)
repeat_region      /rpt_family="L1MA9"
                    complement(19947..20241)
repeat_region      /rpt_family="AluJo"
                    complement(20267..20309)
repeat_region      /rpt_family="Charlie1"
                    complement(20321..20702)
repeat_region      /rpt_family="L1MD2"
                    complement(20713..20824)
repeat_region      /rpt_family="MER5A"
                    complement(21299..21499)
repeat_region      /rpt_family="MIR"
                    complement(21650..21763)
repeat_region      /rpt_family="MIR"
                    complement(21875..21909)
repeat_region      /rpt_family="MER5B"
                    complement(21910..22226)
repeat_region      /rpt_family="AluSx"
                    complement(22227..22484)
repeat_region      /rpt_family="MER5B"
                    complement(22757..23190)
repeat_region      /rpt_family="MLT1J"
                    complement(23470..23775)
repeat_region      /rpt_family="AluSg"
                    complement(23761..23803)
repeat_region      /rpt_family="TTTTA"n"
                    complement(24191..24257)
repeat_region      /rpt_family="(TG)n"
                    complement(25248..25279)
repeat_region      /rpt_family="T-rich"

```

```

Query Match      92.0%; Score 18.4; DB 9; Length 146698;
Best Local Similarity 95.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 GGACATGCCGCGCATGTCC 20
Db      106508 GGACATGCCGCGCATGTCC 106527

RESULT 24
AC021460/c      146698 bp      DNA      linear      PRI 15-MAR-2002
LOCUS      AC021460
DEFINITION      Homo sapiens chromosome 15, clone RP11-259P15, complete sequence.
ACCESSION      AC021460
VERSION      AC021460.10
KEYWORDS      GI:19482349
SOURCE      HTG.
ORGANISM      Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 146698)
AUTHORS      Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckert,R., Beda,F.,
            Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,F.,
            Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenesfor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            JOURNAL      2 (bases 1 to 146698)
            Unpublished
            Homo sapiens chromosome 15, clone RP11-259P15

```

```

REFERENCE
AUTHORS
JOURNAL
REFERENCE
AUTHORS

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS

```

```

TITLE
JOURNAL
COMMENT

```

```

FEATURES
source
Location/Qualifiers
1..146698
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-259P15"
/clone_1lb="RP11-259P15 Human Male BAC"
59..119
/rpt_family="MIR"
repeat_region
216..561
/rpt_family="Aluub"
complement(562..732)
/rpt_family="L2"
1082..1119
/rpt_family="AT-rich"
repeat_region
1120..1415
/rpt_family="Aluub"
complement(1417..1696)

```

```

repeat_region /rpt_family="L1MB8"
complement(1697. .1985)
repeat_region /rpt_family="AluSc"
complement(1986. .2138)
repeat_region /rpt_family="L1MB8"
2798. .2881
repeat_region /rpt_family="MIR"
2958. .3108
repeat_region /rpt_family="LTR33"
complement(3113. .3376)
repeat_region /rpt_family="L1PBA"
3397. .3639
repeat_region /rpt_family="LTR33A"
complement(3640. .3977)
repeat_region /rpt_family="THE1B"
3978. .4097
repeat_region /rpt_family="LTR33A"
complement(4257. .4566)
repeat_region /rpt_family="AluSc"
complement(4971. .5131)
repeat_region /rpt_family="FRAM"
5255. .5494
repeat_region /rpt_family="AluUo"
5649. .5963
repeat_region /rpt_family="AluSc"
6659. .6791
repeat_region /rpt_family="AluUo"
8068. .8244
repeat_region /rpt_family="AluUo"
9323. .9497
repeat_region /rpt_family="L3"
9600. .9871
repeat_region /rpt_family="L3"
9875. .10038
repeat_region /rpt_family="AluUo"
10045. .10340
repeat_region /rpt_family="AluSc"
10371. .10419
repeat_region /rpt_family="AT rich"
complement(10475. .10844)
repeat_region /rpt_family="L2"
complement(11081. .11323)
repeat_region /rpt_family="L1MC4a"
complement(11750. .12055)
repeat_region /rpt_family="AluSc"
12281. .12405
repeat_region /rpt_family="MIR"
12691. .12764
repeat_region /rpt_family="TATATG)n"
13182. .13457
repeat_region /rpt_family="AluSc"
13461. .13484
repeat_region /rpt_family="CAA)n"
13568. .13651
repeat_region /rpt_family="MIR3"
complement(13859. .14002)
repeat_region /rpt_family="MIR3"
complement(14683. .14795)
repeat_region /rpt_family="MIR"
14892. .15185
repeat_region /rpt_family="AluSc"
complement(16265. .16367)
repeat_region /rpt_family="MIR"
16368. .16405
repeat_region /rpt_family="(TG)n"
complement(16406. .16541)
repeat_region /rpt_family="MIR"
complement(16553. .16824)
repeat_region /rpt_family="L1MC4a"
complement(16974. .17273)
repeat_region /rpt_family="AluSc"
complement(17892. .18185)
repeat_region /rpt_family="AluSc"

```

```

repeat_region complement(19150. .19457)
repeat_region /rpt_family="AluU"
complement(19464. .19526)
repeat_region /rpt_family="MER5A"
19527. .19922
repeat_region /rpt_family="L1MA9"
19947. .20241
repeat_region /rpt_family="AluUo"
20267. .20309
repeat_region /rpt_family="Charliel"
20321. .20702
repeat_region /rpt_family="L1MD2"
20713. .20824
repeat_region /rpt_family="MER5A"
complement(21299. .21499)
repeat_region /rpt_family="MIR"
21650. .21763
repeat_region /rpt_family="MIR"
21875. .21909
repeat_region /rpt_family="MER5B"
complement(21910. .22226)
repeat_region /rpt_family="AluSc"
22227. .22484
repeat_region /rpt_family="MER5B"
22757. .23190
repeat_region /rpt_family="MLT1J"
complement(23470. .23775)
repeat_region /rpt_family="AluSc"
23781. .23803
repeat_region /rpt_family="(TTTAA)n"
24191. .24257
repeat_region /rpt_family="(TG)n"
25248. .25279
repeat_region /rpt_family="T-rich"

```

Query Match 92.0%; Score 18.4; DB 9; Length 146698;
 Best Local Similarity 95.0%; Pred. No. 2.3e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 GGACATGCCCGCATGTCC 20
 Db 106527 GGACATGCTGCGCATGTCC 106508

RESULT 25
 ALJ359638 159478 bp DNA linear HTG 10-JUL-2001
 LOCUS Homo sapiens chromosome 6 clone RP11-769A14, 14 unordered pieces.
 DEFINITION
 ACCESSION ALJ359638
 VERSION 7 GI:12539703
 KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Sims,S.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 COMMENT On Jan 26, 2001 this sequence version replaced gi:9864217.
 ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: BA769A14
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 153705 bases at least Q40
 Consensus quality: 155925 bases at least Q30
 Consensus quality: 157184 bases at least Q20
 Insert size: 158178; sum-of-contigs
 Insert size: 154847; agarose-fp
 Quality coverage: 4.26x in Q20 bases; sum-of-contigs Quality
 coverage: 4.51x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 6892: contig of 6892 bp in length
 6893 gap of 100 bp
 6993 contig of 7377 bp in length
 14370 gap of 100 bp
 14469 gap of 100 bp
 19375 contig of 4806 bp in length
 19376 gap of 100 bp
 19476 gap of 100 bp
 33036 contig of 13561 bp in length
 33037 gap of 100 bp
 33137 contig of 8887 bp in length
 42024 gap of 100 bp
 42124 contig of 31779 bp in length
 73903 gap of 100 bp
 74002 gap of 100 bp
 84868 contig of 10866 bp in length
 84869 gap of 100 bp
 84969 contig of 26620 bp in length
 11589 gap of 100 bp
 11688 gap of 100 bp
 131172 contig of 19484 bp in length
 131272 gap of 100 bp
 134785 contig of 3513 bp in length
 134786 gap of 100 bp
 134885 gap of 100 bp
 138106 contig of 3221 bp in length
 138206 gap of 100 bp
 138207 contig of 9534 bp in length
 147740 contig of 9534 bp in length
 147841 gap of 100 bp
 147841 gap of 3829 bp in length
 151669 contig of 3829 bp in length
 151769 gap of 100 bp
 151670 gap of 100 bp
 151770 contig of 7709 bp in length.

FEATURES

source

1. 159478
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="RP11-769A14"
 /clone_1fb="RPC1-11.3"
 1. 6892
 /note="assembly_fragment:00602
 fragment_chain:1"
 6993. 14369
 /note="assembly_fragment:02031
 fragment_chain:1"
 14470. 19375
 /note="assembly_fragment:01886
 fragment_chain:1"
 19476. 33036
 /note="assembly_fragment:01758
 fragment_chain:1"
 33137. 42023
 /note="assembly_fragment:01839
 fragment_chain:1"
 42124. 73902
 /note="assembly_fragment:00383
 fragment_chain:2"
 74003. 84868
 /note="assembly_fragment:01701
 fragment_chain:2"
 84969. 11588

ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 159478;
 Best Local Similarity 95.0%; Pred. No. 2.3e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGACATGCCGGGCATATCC 20
 Db 139362 GGACATGCCGGGCATATCC 139361

RESULT 26

AL359638/c 159478 bp DNA linear HTG 10-JUL-2001
 LOCUS Homo sapiens chromosome 6 clone RP11-769A14, 14 unordered pieces.
 DEFINITION
 ACCESSION AL359638
 VERSION
 KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Sins, S.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jan 26, 2001 this sequence version replaced gi:9864217.

COMMENT

----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: ba769A14
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; 108752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 153705 bases at least Q40
 Consensus quality: 155925 bases at least Q30
 Consensus quality: 157184 bases at least Q20
 Insert size: 158178; sum-of-contigs
 Insert size: 154847; agarose-fp
 Quality coverage: 4.26x in Q20 bases; sum-of-contigs Quality
 coverage: 4.51x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 6892: contig of 6892 bp in length
* 6893 6992: gap of 100 bp
* 14369: contig of 7377 bp in length
* 14370 14469: gap of 100 bp
* 14470 19375: contig of 4906 bp in length
* 19376 19475: gap of 100 bp
* 19476 33036: contig of 13561 bp in length
* 33037 33137: gap of 100 bp
* 33137 42023: contig of 8887 bp in length
* 42024 42124: gap of 100 bp
* 42124 73902: contig of 31779 bp in length
* 73903 74002: gap of 100 bp
* 74003 84868: contig of 10866 bp in length
* 84869 84969: gap of 100 bp
* 84969 111588: contig of 26620 bp in length
* 111589 111689: gap of 100 bp
* 111689 131172: contig of 19484 bp in length
* 131173 131272: gap of 100 bp
* 131273 134785: contig of 3513 bp in length
* 134786 134885: gap of 100 bp
* 134886 138106: contig of 3221 bp in length
* 138107 138206: gap of 100 bp
* 138207 147740: contig of 9534 bp in length
* 147741 147840: gap of 100 bp
* 147841 151693: contig of 3829 bp in length
* 151694 151770: gap of 100 bp
* 151770 159478: contig of 7709 bp in length.

```

FEATURES

source

```

1. 159478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-769A14"
/clone_lib="RPC1-11.3"
1. 6852
/note="assembly_fragment:00602
fragment_chain:1"
6993. 14369
/note="assembly_fragment:02031
fragment_chain:1"
14470. 19375
/note="assembly_fragment:01886
fragment_chain:1"
19476. 33036
/note="assembly_fragment:01758
fragment_chain:1"
33137. 42023
/note="assembly_fragment:01839
fragment_chain:1"
42124. 73902
/note="assembly_fragment:00383
fragment_chain:2"
74003. 84868
/note="assembly_fragment:01701
fragment_chain:2"
84969. 111588
/note="assembly_fragment:00704
fragment_chain:3"
11189. 131172
/note="assembly_fragment:01843
fragment_chain:3"
131273. 134785
/note="assembly_fragment:00807
fragment_chain:4"
134886. 138106
/note="assembly_fragment:01302
fragment_chain:4"
138207. 147740

```

```

misc_feature
/note="assembly_fragment:01109
fragment_chain:5"
147841. 151669
/note="assembly_fragment:01940
fragment_chain:5"
151770. 159478
/note="assembly_fragment:00573
clone_end:T7
vector_side:right"

```

ORIGIN

```

Query Match 92.0%; Score 18.4; DB 2; Length 159478;
Best Local Similarity 95.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Gy 1 GGACATGCCCGGCGATGTCC 20
Db 139381 GGACATGCCCGGCGATGTCC 139362

```

RESULT 27

```

HS44A20 165090 bp DNA linear PRI 04-MAR-2003
LOCUS Human DNA sequence from clone RP1-44A20 on chromosome 6q23.1-24.3,
DEFINITION complete sequence.
ACCESSION AL035086
VERSION AL035086.12 GI:4741478
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 165090)
Matthews, L.
Direct Submission
Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 4, 1999 this sequence version replaced gi:4680389.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk

```

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> RP1-44A20 is from the library RPC1-1 constructed by the group of Pieter de Jong. For further details see <http://www.choyl.org/bacpac/home.htm> VECTOR: pCYPAC2.

FEATURES

Location/Qualifiers

source 1. 165090

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="q23.1-24.3"
/clone="RP1-44A20"
/clone_lib="RPCT-1"

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 165090;
Best Local Similarity 95.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20

Db 67049 GGACATGCCCGGCATGTCC 67068

RESULT 28

LOCUS HS44A20 165090 bp DNA linear PRI 04-MAR-2003
DEFINITION Human DNA sequence from clone RP1-44A20 on chromosome 6q23.1-24.3,
complete sequence.
ACCESSION AL035086
VERSION AL035086.12 GI:4741478
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On May 4, 1999 this sequence version replaced gi:4680389.
JOURNAL Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6
RP1-44A20 is from the library RPCT-1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2

FEATURES
source Location/Qualifiers
1. 165090

/organism="Homo sapiens"

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="q23.1-24.3"
/clone="RP1-44A20"
/clone_lib="RPCT-1"

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 165090;
Best Local Similarity 95.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20

Db 67068 GGACATGCCCGGCATGTCC 67049

RESULT 29

LOCUS AC126323 170132 bp DNA linear PRI 11-NOV-2002
DEFINITION Homo sapiens chromosome 15, clone RP11-299H22, complete sequence.
ACCESSION AC126323
VERSION AC126323.6 GI:24850482
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS Submitted (05-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL Direct Submission
Submitted (05-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 170132)
JOURNAL Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguski, L., Boukhalter, B.,
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunhahang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

REFERENCE

AUTHORS Submitted (05-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL Direct Submission
Submitted (05-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 170132)
JOURNAL Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguski, L., Boukhalter, B.,
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunhahang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguski, L., Boukhalter, B.,
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunhahang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (31-OCT-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 170132)
 Birren, B., Nisbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Baetjen, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
 Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galegan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagob, B., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J.,
 Matthews, C., McCarthy, M., Meldrum, J., Meneses, L., Mihova, T.,
 Mungai, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Notbu, C.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
 Rice, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, K.,
 Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
 Travers, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (11-NOV-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 11, 2002 this sequence version replaced gi:24431891.
 All repeats were identified using RepeatMasker:
 Smt, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
JOURNAL
REFERENCE
AUTHORS

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L27748
 Center clone name: 299_H_22

FEATURES
SOURCE

Location/Qualifiers
 1. 170132
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-299H22"
 /clone_lib="RPC1-11 Human Male BAC"
 complement(92..287)
 repeat_region
 /rpt_family="MIR"
 1956..2179
 /rpt_family="MIR"
 complement(3303..3533)
 repeat_region
 /rpt_family="MIR"
 5010..5166
 /rpt_family="MIR"
 complement(5503..5759)
 repeat_region
 /rpt_family="L2"
 6342..6441
 /rpt_family="L1MD3"
 7226..7270
 /rpt_family="poly(pyrimidine)"
 complement(7275..7354)
 repeat_region
 /rpt_family="LTR16A"
 complement(7491..7674)
 repeat_region
 /rpt_family="L1MCS5"
 7675..7710
 /rpt_family="TTTA)n"
 complement(7711..7782)
 repeat_region
 /rpt_family="L1MCS5"
 complement(7784..7905)
 repeat_region
 /rpt_family="L1MA5A"
 8239..9936
 repeat_region

repeat_region
 9937..10293
 /rpt_family="L1MA4"
 repeat_region
 10295..10897
 /rpt_family="MSTC"
 repeat_region
 10895..10960
 /rpt_family="L1MA4"
 repeat_region
 10961..11197
 /rpt_family="L1MA4"
 repeat_region
 11198..12677
 /rpt_family="L1PB1"
 repeat_region
 12733..12778
 /rpt_family="L1PB1"
 repeat_region
 12779..12968
 /rpt_family="TTA)n"
 complement(12917..12968)
 repeat_region
 13077..13199
 /rpt_family="L3"
 repeat_region
 13201..13593
 /rpt_family="MIR3"
 complement(13593..13747)
 repeat_region
 14902..14923
 /rpt_family="GC_rich"
 15271..15404
 /rpt_family="GC_rich"
 16661..16696
 /rpt_family="U2"
 complement(16882..17005)
 repeat_region
 17405..17491
 /rpt_family="MIR"
 complement(17561..17635)
 repeat_region
 17712..17760
 /rpt_family="MIR"
 repeat_region
 17761..18252
 /rpt_family="CT_rich"
 complement(17975..18252)
 repeat_region
 18253..18688
 /rpt_family="MER2"
 complement(18688..18823)
 repeat_region
 18824..19199
 /rpt_family="MIR"
 complement(19199..19999)
 repeat_region
 20000..20085
 /rpt_family="MIR"
 complement(20085..20168)
 repeat_region
 20169..20449
 /rpt_family="L1ME3A"
 complement(20451..20627)
 repeat_region
 20628..20795
 /rpt_family="AluDb"
 complement(20795..20907)
 repeat_region
 20908..21343
 /rpt_family="L1ME3A"
 complement(21343..21413)
 repeat_region
 21414..21523
 /rpt_family="MER47A"
 complement(21523..21864)
 repeat_region
 21865..21953
 /rpt_family="L3"
 complement(21953..22897)
 repeat_region
 22898..23037
 /rpt_family="MIR"
 complement(23037..23372)
 repeat_region
 23373..23830
 /rpt_family="AluSg"
 complement(23830..25005)
 repeat_region
 25006..25827
 /rpt_family="LTR16A"
 complement(25827..26077)
 repeat_region
 26078..26952
 /rpt_family="MER33"
 complement(26952..27182)
 repeat_region
 27183..27237
 /rpt_family="AT_rich"
 complement(27237..27319)
 repeat_region
 27320..27889
 /rpt_family="MIR"
 complement(27889..27923)
 repeat_region
 27924..28339
 /rpt_family="AT_rich"

repeat_region complement(10961. .11197)
 /rpt_family="L1PB1"
 repeat_region 11198. .12677
 /rpt_family="L1PB1"
 repeat_region 12733. .12778
 /rpt_family="(TA)n"
 repeat_region complement(12917. .12968)
 /rpt_family="L3"
 repeat_region 13077. .13199
 /rpt_family="MIR3"
 repeat_region complement(13593. .13747)
 /rpt_family="MIR"
 repeat_region 14902. .14923
 /rpt_family="GC_rich"
 repeat_region 15271. .15404
 /rpt_family="GC_rich"
 repeat_region 16661. .16696
 /rpt_family="U2"
 repeat_region complement(16882. .17005)
 /rpt_family="MIR"
 repeat_region 17405. .17491
 /rpt_family="L2"
 repeat_region complement(17561. .17635)
 /rpt_family="MIR"
 repeat_region 17712. .17760
 /rpt_family="CT_rich"
 repeat_region complement(17975. .18252)
 /rpt_family="MER2"
 repeat_region complement(18688. .18823)
 /rpt_family="MIR"
 repeat_region complement(18971. .19199)
 /rpt_family="MIR"
 repeat_region complement(19842. .20085)
 /rpt_family="L1M3A"
 repeat_region 20168. .20449
 /rpt_family="MLT1A1"
 repeat_region complement(20451. .20627)
 /rpt_family="AluDb"
 repeat_region 20661. .20795
 /rpt_family="MLT1B"
 repeat_region complement(20804. .20907)
 /rpt_family="L1M3A"
 repeat_region 20979. .21343
 /rpt_family="MERA7A"
 repeat_region 21413. .21523
 /rpt_family="L3"
 repeat_region 21864. .21953
 /rpt_family="MIR"
 repeat_region complement(21977. .22294)
 /rpt_family="AluSg"
 repeat_region 22897. .23037
 /rpt_family="MIR3"
 repeat_region complement(23372. .23830)
 /rpt_family="MLT1C"
 repeat_region complement(24687. .25005)
 /rpt_family="LTR16A"
 repeat_region 25827. .26077
 /rpt_family="MER33"
 repeat_region 26921. .26952
 /rpt_family="AT_rich"
 repeat_region 27182. .27237
 /rpt_family="MIR"
 repeat_region complement(27257. .27319)
 /rpt_family="MIR"
 repeat_region 27889. .27923
 /rpt_family="AT_rich"
 repeat_region 28176. .28351

Query March 92.0% Score 18.4; DB 9; Length 170132;
 Best Local Similarity 95.0%; Pred.No.2.3e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20

Db 113566 GGACATGCCCGGCGCATGTCC 113547
 RESULT 31
 AL670675 173987 bp DNA linear ROD 12-JUL-2002
 LOCUS Mouse DNA sequence from clone RP23-263N18 on chromosome X, complete
 DEFINITION sequence.
 AL670675
 ACCESSION AL670675.7 GI:21531153
 VERSION HTG.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Wray,P.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Jun 21, 2002 this sequence version replaced gi:20330273.
 COMMENT
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquerry@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em., EMBL; Sw.,
 SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
 database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-263N18 is
 from the RPCI-23 Mouse PAC Library
 constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAC3.6.
 Location/Qualifiers
 1. 173987
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="X"
 /clone="RP23-263N18"
 /clone_1bp="RPCI-23"

ORIGIN
 Query March 92.0% Score 18.4; DB 10; Length 173987;
 Best Local Similarity 95.0%; Pred.No.2.3e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20

Db 95248 GGACATGCCCGGCGCATGTCC 95267
 RESULT 32
 AL670675 173987 bp DNA linear ROD 12-JUL-2002
 LOCUS

DEFINITION Mouse DNA sequence from clone RP23-263N18 on chromosome X, complete sequence.

ACCESSION AL670675 GI:21531153

VERSION AL670675.7

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 173987)

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE Bases 1 to 173987

JOURNAL Direct Submission

COMMENT Submitted (12-JUL-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jun 21, 2002 this sequence version replaced gi:20330273.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-263N18 is from the RP23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

Location/Qualifiers

1..173987

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="X"

/clone="RP23-263N18"

/clone_11b="RP23-23"

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 173987;

Best Local Similarity 95.0%; Pred. No. 2.3e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGACATGCCCGGCATGTC 20

|||||

Db 95267 GGACCTGCCCGGCATCTCC 95248

RESULT 33

AL773522/c

LOCUS AL773522

DEFINITION Mouse DNA sequence from clone RP23-148P20 on chromosome 4, complete sequence.

ACCESSION AL773522

VERSION AL773522.15 GI:28564418

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 185592)

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE Bases 1 to 185592

JOURNAL Direct Submission

COMMENT Submitted (25-FEB-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Feb 25, 2003 this sequence version replaced gi:28268662. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-148P20 is from the RP23 Mouse BAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

Location/Qualifiers

1..185592

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="4"

/clone="RP23-148P20"

/clone_11b="RP23-23"

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 185592;

Best Local Similarity 95.0%; Pred. No. 2.2e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGACATGCCCGGCATGTC 20

|||||

Db 140068 GGACATGCCCGGCATGCC 140067

RESULT 34

AL773522/c

LOCUS AL773522

DEFINITION Mouse DNA sequence from clone RP23-148P20 on chromosome 4, complete sequence.

ACCESSION AL773522

VERSION AL773522.15 GI:28564418

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 185592)
 AUTHORS Phillimore,B.
 TITLE Direct Submission
 JOURNAL Submitted (25-FEB-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 COMMENT On Feb 25, 2003 this sequence version replaced gi:28268662. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

 Center: Wellcome Trust Sanger Institute
 Genome Center
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EMBL, SW, SWISSPROT, Tr, TREMBL, Wp, WORMPEP, Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 from the RPCI-23 Mouse BAC Library
 constructed by the group of Pieter de Jong.
 For further details see http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6.
 Location/Qualifiers
 1..185592
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="4"
 /clone="RP23-148P20"
 /clone_11b="RPCI-23"
 /clone_11b="RPCI-23"
 ORIGIN
 Query Match 92.0%; Score 18.4; DB 10; Length 185592;
 Best Local Similarity 95.0%; Pred. No. 2.2e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GGACATGCCGGGCGATGTC 20
 Db 140087 GGGCATGCCGGGCGATGTC 140068
 RESULT 35
 AC139327
 LOCUS Mus musculus chromosome UNK clone RP24-362N3, WORKING DRAFT
 DEFINITION
 SEQUENCE, 5 unordered pieces.
 AC139327
 AC139327.2 GI:28173253
 VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FUTLTOP.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 191128)
 McPherson,J.D. and Waterston,R.H.

TITLE The sequence of Mus musculus clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 191128)
 AUTHORS McPherson,J.D. and Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (30-JUN-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 COMMENT On Jan 31, 2003 this sequence version replaced gi:28144456.

 Center: Washington University Genome Sequencing Center
 Genome Center
 Center code: WUSGC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Contact: submissions@watson.wustl.edu
 Project Information
 Center project name: M_BB0362N03

 Summary Statistics
 Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-primer ET; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 190351 bases at least Q40
 Consensus quality: 190555 bases at least Q30
 Consensus quality: 190792 bases at least Q20
 Insert size: 164000; agarose-fp
 Insert size: 193202; sum-of-contigs
 Quality coverage: 13.72 in Q20 bases; agarose-fp
 Quality coverage: 12.92 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1..8971: contig of 8971 bp in length
 8972..9071: gap of unknown length
 9072..26264: contig of 17193 bp in length
 26265..26364: gap of unknown length
 26365..59759: contig of 33395 bp in length
 59760..59859: gap of unknown length
 59860..101845: contig of 41986 bp in length
 101846..101945: gap of unknown length
 101946..191128: contig of 89183 bp in length.
 Location/Qualifiers
 1..191128
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="UNK"
 /clone="RP24-362N3"
 1..8971
 /note="assembly_name:Contig47"
 9072..26264
 /note="assembly_name:Contig48"
 26365..59759
 /note="assembly_name:Contig49"
 59860..101845
 /note="assembly_name:Contig50"
 101946..191128
 /note="assembly_name:Contig51"
 ORIGIN
 Query Match 92.0%; Score 18.4; DB 2; Length 191128;

Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCATGTC 20

Db 99262 GGACCTGCCCGGCATGTC 99281

RESULT 36

AC139327/c

LOCUS AC139327 191128 bp DNA linear HTG 21-FEB-2003
DEFINITION Mus musculus chromosome UNK clone RP24-362N3, WORKING DRAFT
SEQUENCE, 5 unordered pieces.

ACCESSION

AC139327

VERSION

AC139327.2

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 191128)

AUTHORS

McPherson, J.D. and Waterston, R.H.

TITLE

The sequence of Mus musculus clone

REFERENCE

2 (bases 1 to 191128)

AUTHORS

McPherson, J.D. and Waterston, R.H.

TITLE

Unpublished

REFERENCE

Submitted (30-JUN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

AUTHORS

McPherson, J.D. and Waterston, R.H.

TITLE

Direct Submission

REFERENCE

Submitted (21-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

AUTHORS

On Jan 31, 2003 this sequence version replaced gi:28144456.

TITLE

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

COMMENT

On Jan 31, 2003 this sequence version replaced gi:28144456.

* 101846 101945: gap of unknown length
* 101946 191128: contig of 89183 bp in length.
Location/Qualifiers
1. 191128

FEATURES

Source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

* 101846 101945: gap of unknown length
* 101946 191128: contig of 89183 bp in length.
Location/Qualifiers
1. 191128

FEATURES

Source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

* NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 1412: contig of 1412 bp in length
1413 1512: gap of unknown length
1513 3635: contig of 2123 bp in length
3636 3735: gap of unknown length
3736 4999: contig of 1264 bp in length
5000 5099: gap of unknown length
5100 7667: contig of 2568 bp in length
7668 7767: gap of unknown length
7768 10116: contig of 2349 bp in length
10117 10217: gap of unknown length
10218 14002: contig of 3786 bp in length
14003 14102: gap of unknown length
14103 17164: contig of 3062 bp in length
17165 17264: gap of unknown length
17265 20585: contig of 3321 bp in length
20586 20685: gap of unknown length
20686 24370: contig of 3685 bp in length
24371 24471: gap of unknown length
24472 29951: contig of 5481 bp in length
29952 30051: gap of unknown length
30052 37837: contig of 7786 bp in length
37838 37937: gap of unknown length
37938 48151: contig of 10214 bp in length
48152 48251: gap of unknown length
48252 61495: contig of 13244 bp in length
61496 61595: gap of unknown length
61596 73344: contig of 11749 bp in length
73345 73444: gap of unknown length
73445 85338: contig of 11894 bp in length
85339 85439: gap of unknown length
85440 97336: contig of 11898 bp in length
97337 97436: gap of unknown length
97437 109646: contig of 12210 bp in length
109647 109747: gap of unknown length
109748 123544: contig of 13798 bp in length
123545 123644: gap of unknown length
123645 139857: contig of 16213 bp in length
139858 139957: gap of unknown length
139959 154527: contig of 14570 bp in length
154528 154627: gap of unknown length
154628 174231: contig of 19604 bp in length
174232 174331: gap of unknown length
174332 202227: contig of 27896 bp in length.
174332
```

FEATURES
source
1. .202227
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/clone="RP11-546114"
1. .1412
/note="assembly_name:Contig6"
1513. .3635
/note="assembly_name:Contig7"
3736. .4999
/note="assembly_name:Contig8"
5100. .7667
/note="assembly_name:Contig9"
7768. .10116
/note="assembly_name:Contig10"
10217. .14002
/note="assembly_name:Contig11"
14103. .17164
/note="assembly_name:Contig12

```
clone end:17  
vector_side:left"  
17265..20585  
/note="assembly_name:Contig13"  
20586..24370  
/note="assembly_name:Contig14"  
24471..29951  
/note="assembly_name:Contig15"  
30052..37837  
/note="assembly_name:Contig16"  
37938..48151  
/note="assembly_name:Contig17"  
48252..61495  
/note="assembly_name:Contig18"  
61596..73344  
/note="assembly_name:Contig19"  
73445..85338  
/note="assembly_name:Contig20  
clone end:SP6  
vector_side:right"  
85439..97336  
/note="assembly_name:Contig21"  
97437..109646  
/note="assembly_name:Contig22"  
109747..123544  
/note="assembly_name:Contig23"  
123645..139857  
/note="assembly_name:Contig24"  
139958..154527  
/note="assembly_name:Contig25"  
154628..174231  
/note="assembly_name:Contig26"  
174332..202227  
/note="assembly_name:Contig27"
```

ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 202227;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

0y 1 GGACATGCGCGGCGATGTC 20

Db 184949 GGACATGCGCTGGCGATGTC 184968

RESULT 38

AC068126/c

LOCUS AC068126 202227 bp DNA linear HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 15 clone RP11-546114, WORKING DRAFT
SEQUENCE, 22 unordered pieces.

ACCESSION AC068126.2 GI:7705251

VERSION AC068126.2 HTG: HTGS PHASE1; HTGS_DRAFT.
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORANISM Homo sapiens

REFERENCE

REFERENCE 1 (bases 1 to 202227)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 202227)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

COMMENT On May 4, 2000 this sequence version replaced gi:7658489.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml

```

----- Project Information -----
Center project name: H_NH0546114
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-Primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 188247 bases at least Q40
Consensus quality: 193018 bases at least Q30
Insert size: 174000; agarose-fp
Insert size: 200127; sum-of-contigs
Quality coverage: 4.45 in Q20 bases; agarose-fp
Quality coverage: 3.93 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1412: contig of 1412 bp in length
1413
1512: gap of unknown length
1513
3635: contig of 2123 bp in length
3735: gap of unknown length
3736
4999: contig of 1264 bp in length
5099: gap of unknown length
5100
7667: contig of 2568 bp in length
7668
7768
10116: contig of 2349 bp in length
10117
10216: gap of unknown length
14002: contig of 3786 bp in length
14003
14102: gap of unknown length
14103
17164: contig of 3062 bp in length
17165
17264: gap of unknown length
17265
20585: contig of 3321 bp in length
20586
20685: gap of unknown length
24370: contig of 3685 bp in length
24371
24470: gap of unknown length
29951: contig of 5481 bp in length
29952
30051: gap of unknown length
37837: contig of 7786 bp in length
37838
37937: gap of unknown length
48151: contig of 10214 bp in length
48152
48251: gap of unknown length
61495: contig of 13244 bp in length
61496
61595: gap of unknown length
73444: contig of 11749 bp in length
73445
73446: gap of unknown length
85338: contig of 11894 bp in length
85339
85438: gap of unknown length
85439
97336: contig of 11898 bp in length
97337
97436: gap of unknown length
109646: contig of 12210 bp in length
109647
109746: gap of unknown length
123544: contig of 13798 bp in length
123545
123644: gap of unknown length
123645
139857: contig of 16213 bp in length
139858
139957: gap of unknown length
139958
14527: contig of 14570 bp in length
14528
154627: gap of unknown length
154628
174231: contig of 19604 bp in length
174232
174331: gap of unknown length
174332
202227: contig of 27836 bp in length.
Location/Qualifiers
1. 202227
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"

```

```

misc_feature
/clone="RP11-546114"
1. 1412
/note="assembly_name:Contig6"
1513. 3635
/note="assembly_name:Contig7"
3736. 4999
/note="assembly_name:Contig8"
5100. 7667
/note="assembly_name:Contig9"
7768. 10116
/note="assembly_name:Contig10"
10217. 14002
/note="assembly_name:Contig11"
14103. 17164
/note="assembly_name:Contig12
clone_end:T7
vector_side:left"
17265. 20585
/note="assembly_name:Contig13"
20686. 24370
/note="assembly_name:Contig14"
24471. 29951
/note="assembly_name:Contig15"
30052. 37837
/note="assembly_name:Contig16"
37938. 48151
/note="assembly_name:Contig17"
48252. 61495
/note="assembly_name:Contig18"
61596. 73344
/note="assembly_name:Contig19"
73445. 85338
/note="assembly_name:Contig20
clone_end:SP6
vector_side:right"
85439. 97336
/note="assembly_name:Contig21"
97437. 109646
/note="assembly_name:Contig22"
109747. 123544
/note="assembly_name:Contig23"
123645. 139857
/note="assembly_name:Contig24"
139958. 154527
/note="assembly_name:Contig25"
154628. 174231
/note="assembly_name:Contig26"
174332. 202227
/note="assembly_name:Contig27"
-----
ORIGIN
Query Match 92.0%; Score 18.4; DB 2; Length 202227;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGACATGCCCGGACATGTC 20
Db 184968 GGACATGCCCGGACATGTC 184949
-----
RESULT 39
AC087483 203950 bp DNA linear HTG 24-MAY-2001
LOCUS AC087483
DEFINITION Homo sapiens chromosome 15 clone RP11-546114 map 15, WORKING DRAFT
SEQUENCE, 6 unordered pieces.
ACCESSION AC087483
VERSION AC087483.3 GI:14192959
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 203950)

```

AUTHORS TITLE JOURNAL REFERENCE AUTHORS

Bitren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone RP11-546114
Unpublished
2 (bases 1 to 203950)

Bitren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Baetien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
Camarata,A., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., Lacroque,K., Lamazares,R., Landers,T.,
Lecoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
McChetters,N., Mathews,C., McCarthy,M., McGowan,P., McKernan,K.,
McChetters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunhkhang,P., Piere,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnaz,C., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,
Travers,M., Travis,N., Trigglio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE JOURNAL COMMENT

Submitted (05-JUN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 24, 2001 this sequence version replaced g1:13470198.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W1BR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L11941

Center clone name: 546_1_14

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 202418 bases at least Q40

Consensus quality: 203020 bases at least Q30

Consensus quality: 203224 bases at least Q20

Insert size: 205000; agarose-fp

Insert size: 203450; sum-of-contrigs

Quality coverage: 9.5 in Q20 bases; agarose-fp

Quality coverage: 9.6 in Q20 bases; sum-of-contrigs

NOTE: This is a 'working draft' sequence. It currently
consists of 6 contrigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contrigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1701: contrig of 1701 bp in length
* 1702 1801: gap of 100 bp
* 1802 17309: contrig of 15508 bp in length
* 17310 17409: gap of 100 bp
* 17410 60846: contrig of 43437 bp in length
* 60847 60946: gap of 100 bp
* 60947 108903: contrig of 47957 bp in length
* 108904 109003: gap of 100 bp
* 109004 158992: contrig of 49989 bp in length
* 158993 159092: gap of 100 bp
* 159093 203950: contrig of 44858 bp in length.

FEATURES SOURCE

Location/Qualifiers
1. 203950

ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 203950;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Ct 1 GCACATGCCCGGCGCATGCC 20
Db 163747 GCACATGCCCGGCGCATGCC 163766

RESULT 40 LOCUS DEFINITION

AC087483 203950 bp DNA linear HTG 24-MAY-2001
Homo sapiens chromosome 15 clone RP11-546114 map 15, WORKING DRAFT
SEQUENCE, 6 unordered pieces.

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AC087483
AC087483.3 GI:14192959
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Homo sapiens (human)
Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-546114"
/clone_1lb="RPC1-11 Human Male BAC"
1..1701
/note="assembly_fragment
clone_end:sp6
vector_side:left"
1802..17309
/note="assembly_fragment"
17410..60846
/note="assembly_fragment"
60947..108903
/note="assembly_fragment"
109004..158992
/note="assembly_fragment"
159093..203950
/note="assembly_fragment
clone_end:t7
vector_side:right"

JOURNAL
Submitted (05-JUN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 24, 2001 this sequence version replaced gi:13470198.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: 111941

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 202418 bases at least Q40

Consensus quality: 203020 bases at least Q30

Consensus quality: 203224 bases at least Q20

Insert size: 205000; agarose-fp

Insert size: 203450; sum-of-contigs

Quality coverage: 9.5 in Q20 bases; agarose-fp

Quality coverage: 9.6 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

FEATURES
source
1. 1701: contig of 1701 bp in length
1702 1801: gap of 100 bp
1802 17309: contig of 15508 bp in length
17310 17409: gap of 100 bp
17410 60846: contig of 43437 bp in length
60847 60946: gap of 100 bp
60947 108903: contig of 47957 bp in length
108904 109003: gap of 100 bp
109004 158992: contig of 49989 bp in length
158993 159092: gap of 100 bp
159093 203950: contig of 44858 bp in length.

Location/Qualifiers
1. 203950
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"

/clone="RP11-546114"
/clone_lib="RPCT-11 Human Male BAC"
1. 1701

/note="assembly_fragment
clone_end:SP6
vector_side:left"

misc_feature
1802. 17309
/note="assembly_fragment"

misc_feature
17410. 60846
/note="assembly_fragment"

misc_feature
60947. 108903
/note="assembly_fragment"

misc_feature
109004. 158992
/note="assembly_fragment"

misc_feature
159093. 203950
/note="assembly_fragment"

misc_feature
clone_end:T7
vector_side:right"

ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 203950;

Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Ct 1 GGACATGCCCGGACATGTC 20
Db 163766 GGACATGCTGGCATGTCC 163747

RESULT 41

AC132452 218688 bp DNA linear ROD 25-NOV-2003
LOCUS Mus musculus BAC clone RP23-195K10 from chromosome 19, complete
DEFINITION sequence.

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

AC132452.3 GI:33005055
AC132452.3

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., paired quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

COMMENT

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tatoro in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC133523.

Location/Qualifiers

```

1. 218688
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="19"
/map="19"
/clone="RP23-195K10"
/clone_id="RPCI-23"
291..427
/rpt_family="Alu"
repeat_region
444..629
/rpt_family="B2"
repeat_region
769..808
/rpt_family="B2"
repeat_region
6444..6561
/rpt_family="B4"
repeat_region
7006..7149
/rpt_family="Alu"
repeat_region
7423..7554
/rpt_family="Alu"
repeat_region
10624..10683
/rpt_family="B4"
repeat_region
12467..12553
/rpt_family="B2"
repeat_region
12472..12544
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=37.16 / Sec struct
Sc=-11.21)"
13455..13544
/rpt_family="B2"
repeat_region
13789..13923
/rpt_family="Alu"
repeat_region
13942..14096
/rpt_family="B2"
repeat_region
14342..14460
/rpt_family="B4"
repeat_region
14594..14703
/rpt_family="Alu"
repeat_region
16272..16393
/rpt_family="MERL_type"
repeat_region
16796..16882
/rpt_family="MALR"
repeat_region
16883..17009
/rpt_family="Alu"
repeat_region
17738..17807
/rpt_family="ID"
repeat_region
17858..17964
/rpt_family="Alu"
repeat_region
18293..18501
/rpt_family="B2"
repeat_region
18472..18537
/rpt_family="B2"
repeat_region
18507..18573
/rpt_family="B2"
repeat_region
18543..18609
/rpt_family="B2"
repeat_region
18579..18645
repeat_region

```

```

/rpt_family="B2"
repeat_region
18615..18695
/rpt_family="B2"
repeat_region
18922..19153
/rpt_family="B4"
repeat_region
19341..19482
/rpt_family="Alu"
repeat_region
19711..19845
/rpt_family="Alu"
repeat_region
20067..20185
/rpt_family="Alu"
repeat_region
20277..20423
/rpt_family="Alu"
repeat_region
20540..20582
/rpt_family="B4"
repeat_region
20708..20824
/rpt_family="B4"
repeat_region
21019..21085
/rpt_family="Alu"
repeat_region
21225..21271
/rpt_family="ERV1"
repeat_region
21272..21410
/rpt_family="B4"
repeat_region
21720..21862
/rpt_family="MALR"
repeat_region
21762..21892
/rpt_family="Alu"
repeat_region
22073..22168
/rpt_family="Alu"
repeat_region
22242..22348
/rpt_family="B4"
repeat_region
22379..22497
/rpt_family="Alu"
repeat_region
22819..22962
/rpt_family="Alu"
repeat_region
23189..23234
/rpt_family="Alu"
repeat_region
23290..23463
/rpt_family="B2"
repeat_region
24095..24243
/rpt_family="Alu"
repeat_region
24305..24414
/rpt_family="Alu"
repeat_region
24449..24631
/rpt_family="B2"
repeat_region
24454..24524
/product="tRNA-Glu"
/note="Likely pseudogene (HMM Sc=23.47 / Sec struct
Sc=0.04)"
24838..24947
/rpt_family="Alu"
repeat_region
25090..25148
/rpt_family="ID"
repeat_region
26564..26610
/rpt_family="Alu"
repeat_region
28077..28198
/rpt_family="B2"
repeat_region
28235..28372
/rpt_family="Alu"
repeat_region
28391..28447
/rpt_family="B2"
repeat_region
28476..28657
/rpt_family="B2"
repeat_region
28481..28552
tRNA

```

Query Match 92.0%; Score 18.4; DB 10; Length 218688;
 Best Local Similarity 95.0%; Pred. No. 2.2e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGATGTC 20
 |||||
 DB 60291 GGACATGCCCGGATGTC 60310


```

repeat_region      19711..19845
                    /rpt_family="Alu"
repeat_region      20667..20185
                    /rpt_family="Alu"
repeat_region      20277..20423
                    /rpt_family="Alu"
repeat_region      20540..20582
                    /rpt_family="B4"
repeat_region      20708..20824
                    /rpt_family="B4"
repeat_region      21019..21085
                    /rpt_family="Alu"
repeat_region      21225..21271
                    /rpt_family="ERV1"
repeat_region      21272..21410
                    /rpt_family="B4"
repeat_region      21720..21862
                    /rpt_family="MALR"
repeat_region      21762..21892
                    /rpt_family="Alu"
repeat_region      22073..22168
                    /rpt_family="Alu"
repeat_region      22242..22348
                    /rpt_family="B4"
repeat_region      22379..22497
                    /rpt_family="Alu"
repeat_region      22819..22962
                    /rpt_family="Alu"
repeat_region      23189..23234
                    /rpt_family="Alu"
repeat_region      23290..23463
                    /rpt_family="B2"
repeat_region      24095..24243
                    /rpt_family="Alu"
repeat_region      24305..24414
                    /rpt_family="Alu"
repeat_region      24449..24631
                    /rpt_family="B2"
repeat_region      24454..24524
                    /product="tRNA-Glu"
                    /note="Likely pseudogene (MM Sc=23.47 / Sec struct
                    Sc=0.04)"
repeat_region      24838..24947
                    /rpt_family="Alu"
repeat_region      25090..25148
                    /rpt_family="ID"
repeat_region      26564..26610
                    /rpt_family="Alu"
repeat_region      28077..28198
                    /rpt_family="B2"
repeat_region      28335..28372
                    /rpt_family="Alu"
repeat_region      28391..28447
                    /rpt_family="B2"
repeat_region      28476..28657
                    /rpt_family="B2"
repeat_region      28481..28552
                    /rpt_family="B2"
tRNA

```

Query Match 92.0%; Score 18.4; DB 10; Length 218688;
 Best Local Similarity 95.0%; Pred. No. 2.2e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GGACATGCCGGGATGTC 20
 60310 GGACATGCCGGGATGTC 60291

RESULT 43
 AC113311 260120 bp DNA linear ROD 29-JAN-2004
 LOCUS AC113311 Mus musculus chromosome 1, clone RP23-440J24, complete sequence.
 AC113311
 AC113311.9 GI:41386878

KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

HTG.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 260120)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus chromosome 1, clone RP23-440J24
 Unpublished
 2 (bases 1 to 260120)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., Dekrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S.,
 Ferrelira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Glinde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hago,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Labocque,K., Lamazares,R.,
 Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
 Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M.,
 McEwan,P., McKernan,K., Meldrim,J., Menes,L., Mihova,T.,
 Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
 Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
 Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Stojanovic,N.,
 Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Straus,N., Subramanian,A., Talamas,J., Teffaye,S., Theodore,J.,
 Tophan,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,M., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (28-FEB-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 260120)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
 Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
 Collymore,A., Cook,A., Cooke,P., Corum,B., Dekrellano,K.,
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S.,
 Ferrelira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hago,B.,
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
 Meldrim,J., Menes,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbins,M.,
 Talamas,J., Teffaye,S., Theodore,J., Tophan,K., Travers,M.,
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (17-DEC-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 260120)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
 Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
 Collymore,A., Cook,A., Cooke,P., Corum,B., Dekrellano,K.,
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S.,
 Ferrelira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hago,B.,
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
 Meldrim,J., Menes,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbins,M.,
 Talamas,J., Teffaye,S., Theodore,J., Tophan,K., Travers,M.,
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (17-DEC-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

TITLE
JOURNAL

COMMENT

O'Neil, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talama, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vasilev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (29-JAN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 29, 2004 this sequence version replaced gi:3979533.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBK
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L22773
Center clone name: 440_J_24

FEATURES

source

misc_feature

Location/Qualifiers
1..260120
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP23-440J24"
/clone_id="RPC1-23 Female Mouse BAC"
1..37846
/note="wgs end extension
clone end:SP6"
repeat_region
131..178
/rpt_family="(TG)n"
repeat_region
180..313
/rpt_family="(L2)n"
repeat_region
974..1253
/rpt_family="(RMR6B)
complement(11327..1727)
repeat_region
1733..1934
/rpt_family="(Lx5"
repeat_region
2563..2598
/rpt_family="(RMR6B"
complement(2599..2742)
repeat_region
3780..3810
/rpt_family="(T)n"
complement(3891..4578)
repeat_region
4606..4690
/rpt_family="(L1MA6"
repeat_region
4935..4958
/rpt_family="(B4A"
repeat_region
5026..5089
/rpt_family="(AT_rich"
repeat_region
5702..6084
/rpt_family="(MTD"
complement(6304..6426)
repeat_region
6488..6589
/rpt_family="(MIR"
complement(6964..7174)
repeat_region
7206..7263
/rpt_family="(7SK"
repeat_region
7849..7873
/rpt_family="(Charl1ela"
/rpt_family="(TG)n"

repeat_region 7873..7908
/rpt_family="(GA)n"
complement(9173..9698)
repeat_region
/rpt_family="(Lx"
complement(9791..9848)
repeat_region
/rpt_family="(B3"
9971..10126
/rpt_family="(MIR"
complement(10237..10433)
repeat_region
/rpt_family="(B3A"
10849..11080
/rpt_family="(B4"
complement(11170..11371)
repeat_region
/rpt_family="(CA)n"
complement(11515..11626)
repeat_region
/rpt_family="(Looper"
11713..11743
complement(11922..12107)
repeat_region
/rpt_family="(Looper"
12108..12238
/rpt_family="(RSINE1"
12239..12275
/rpt_family="(TG)n"
repeat_region
12686..12738
/rpt_family="(AT_rich"
13220..13240
/rpt_family="(AT_rich"
13511..13538
/rpt_family="(AT_rich"
13548..13696
/rpt_family="(RSINE1"
14736..14757
/rpt_family="(AT_rich"
complement(14930..14998)
/note="single clone coverage"
complement(15003..15085)
repeat_region
/rpt_family="(PB1D10"
complement(15450..15640)
repeat_region
/rpt_family="(B2_Mm2"
complement(15670..15904)
repeat_region
/rpt_family="(L1"
16334..16373
/rpt_family="(AT_rich"
complement(16493..16674)
repeat_region
/rpt_family="(ORR1D"
16990..17154
/rpt_family="(URR1A"
18058..18109
repeat_region
/rpt_family="(AT_rich"

Query Match

Best Local Similarity 95.0%; Score 18.4; DB 10; Length 260120;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 81584 GGACATGCCCGGCGCATGTCC 81603

RESULT 44
AC113311 260120 bp DNA linear ROD 29-JAN-2004
LOCUS Mus musculus chromosome 1, clone RP23-440J24, complete sequence.
DEFINITION AC113311
ACCESSION AC113311.9 GI:41386678
VERSION
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 260120)
Birren, B., Nusbaum, C. and Lander, B.

TITLE
JOURNAL
REFERENCE
AUTHORS

Mus musculus chromosome 1, clone RP23-440J24
Unpublished
(bases 1 to 260120)

Birtren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barina,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kelle,C., Laroque,K., Lamacares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,Q., Maclean,C.,
Mackdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McGowan,P., McKernan,K., Meldrum,J., Menus,L., Minova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (28-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
(bases 1 to 260120)

Birtren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barina,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepe,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kelle,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
Mackdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Menus,L., Minova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smtt,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (17-DEC-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
(bases 1 to 260120)

Birtren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barina,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepe,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kelle,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
Mackdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Menus,L., Minova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smtt,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (29-JAN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 29, 2004 this sequence version replaced gi:3979533.
All repeats were identified using RepeatMasker:
Smtt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIRR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L22773
Center clone name: 440_J_24

FEATURES
source

location/Qualifiers
1..260120
/organism="Mus musculus"
/mol_type="Genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP23-440J24"
/clone_id="RP23-440J24 Female Mouse BAC"
1..37646
/note="wgs_end_extension
clone_end:5p6"
131..178
/rpt_family="TG)n"
180..313
/rpt_family="L2"
974..1253
/rpt_family="RMER6B"
complement(1327..1727)
/rpt_family="Lx5"
1733..1934
/rpt_family="RMER6B"
2563..2598
/rpt_family="(TTTG)n"
complement(2599..2742)
/rpt_family="B1 MM"
complement(2940..3147)
/rpt_family="L1MA4A"
3780..3810
/rpt_family="(T)n"
complement(3891..4578)
/rpt_family="L1MA6"
4606..4690
/rpt_family="B4A"
4935..4958
/rpt_family="AT_rich"
5026..5089
/rpt_family="AT_rich"
5702..6084
/rpt_family="MTD"
complement(6304..6426)
/rpt_family="MIR"
6498..6589
/rpt_family="Char161a"
complement(6964..7174)
/rpt_family="7SK"
7206..7263
/rpt_family="Char161a"
7849..7873
/rpt_family="(TG)n"
7873..7908
/rpt_family="(GA)n"
complement(9173..9698)
/rpt_family="Lx"
complement(9791..9848)
/rpt_family="B3"
9971..10126
repeat_region

```

repeat_region /rpt_family="MIR"
complement(10237..10433)
/rpt_family="B3A"
repeat_region 10849..11080
/rpt_family="B4"
complement(11170..11371)
repeat_region /rpt_family="looper"
complement(11515..11626)
/rpt_family="looper"
repeat_region 11713..11743
/rpt_family="CA)n"
complement(11922..12107)
repeat_region /rpt_family="looper"
12108..12238
/rpt_family="RSINE1"
repeat_region 12239..12275
/rpt_family="TGG)n"
12686..12738
/rpt_family="AT_rich"
repeat_region 13220..13240
/rpt_family="AT_rich"
repeat_region 13511..13538
/rpt_family="AT_rich"
repeat_region 13548..13696
/rpt_family="RSINE1"
repeat_region 14736..14757
/rpt_family="AT_rich"
complement(14930..14998)
/note="single clone coverage"
repeat_region complement(15003..15085)
/rpt_family="PB1D10"
complement(15450..15640)
/rpt_family="B2_Mm2"
repeat_region complement(15670..15904)
/rpt_family="L1"
repeat_region 16334..16373
/rpt_family="AT_rich"
repeat_region complement(16493..16674)
/rpt_family="ORR1D"
repeat_region 16980..17154
/rpt_family="ORR1A"
18058..18109
/rpt_family="AT_rich"

```

Query Match 92.0%; Score 18.4; DB 10; Length 260120;
 Best Local Similarity 95.0%; Pred. No. 2.2e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
 DB 81603 GGACCTGCGCGGCGCATGTCC 81584

RESULT 45
 AL807771 233077 bp DNA linear ROD 26-NOV-2002
 LOCUS Mouse DNA sequence from clone RP23-141C15 on chromosome 4, complete
 DEFINITION sequence.
 ACCESSION AL807771
 VERSION AL807771.6 GI:25809489
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 233077)
 AUTHORS Lovell, J.
 TITLE Direct Submission
 JOURNAL Submitted (26-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Nov 27, 2002 this sequence version replaced gi:21912779.
 COMMENT ----- Genome Center

Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-141C15 is from the RP23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: PBAC3.6

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

FEATURES

source
 1..233077
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="4"
 /clone="RP23-141C15"
 /clone_1bp="RPC1-23"

ORIGIN

Query Match 90.0%; Score 18; DB 10; Length 233077;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGT 18
 DB 193609 GGACATGCCCGGCGCATGT 193626

RESULT 46
 AL807771/c 233077 bp DNA linear ROD 26-NOV-2002
 LOCUS Mouse DNA sequence from clone RP23-141C15 on chromosome 4, complete
 DEFINITION sequence.
 ACCESSION AL807771
 VERSION AL807771.6 GI:25809489
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 233077)
 AUTHORS Lovell, J.
 TITLE Direct Submission
 JOURNAL Submitted (26-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Nov 27, 2002 this sequence version replaced gi:21912779.
 COMMENT ----- Genome Center
 Center code: SC

Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-141C15 is from the RP23 Mouse PAC library

constructed by the group of Pieter de Jong.

For further details see <http://www.choiti.org/bacpac/home.htm>

VECTOR: pBAC3.6

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

Sequencing from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

FEATURES

source

1. 233077
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-141C15"
/clone_1fb="RP23-141C15"
/clone_2fb="RP23-141C15"

ORIGIN

Query Match 90.0%; Score 18; DB 10; Length 233077;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ACATGCGCGGCGATGTC 20
Db 193626 ACATGCGCGGCGATGTC 193609

RESULT 47

AC093466

LOCUS AC093466 263507 bp DNA linear HTG 23-APR-2003
DEFINITION Mus musculus clone RP23-4M7, WORKING DRAFT SEQUENCE, 11 unordered
pieces.

AC093466

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

2 (bases 1 to 263507)
Unpublished
Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgibbon, W., Gage, D., Galagan, J., Gardyna, S., Gande, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamaz, C., Lander, E., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margis, N., Matthews, C., McCarthy, M., McMan, P., McKernan, K., McPherson, R., Melton, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Pollard, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tophan, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (26-APR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 263507)
Birren, B., Nussbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Cornu, B., Deatellano, K., Diaz, J. S., Dodge, S., Dooly, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgibbon, W., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lander, E., Levine, R., Lindblad-Toh, K., Liu, G., Liu, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melton, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, K., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 17, 2003 this sequence version replaced gs:15290913.
All repeats were identified using RepeatMasker:
Smit, A. P. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13364
Center clone name: 4_M_7

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 50429: contig of 50429 bp in length
* 50430 50529: gap of 100 bp
* 50530 52201: contig of 1672 bp in length
* 52202 52301: gap of 100 bp
* 52302 54425: contig of 2124 bp in length
* 54426 54525: gap of 100 bp
* 54526 57746: contig of 3221 bp in length
* 57747 57847: gap of 100 bp
* 57847 60652: contig of 2806 bp in length
* 60653 60752: gap of 100 bp

```

* 60753 64812: contig of 4060 bp in length
* 64813 64912: gap of 100 bp
* 64913 79944: contig of 15032 bp in length
* 79945 80044: gap of 100 bp
* 80045 101704: contig of 21660 bp in length
* 101705 101804: gap of 100 bp
* 101805 169526: contig of 67722 bp in length
* 169527 203765: gap of 100 bp
* 203766 203866: contig of 34139 bp in length
* 203867 263507: contig of 100 bp
* 263508 263507: contig of 59642 bp in length.

```

FEATURES

source

```

1. 263507
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-4M7"
/clone_lib="RP23-4M7"
/clone_id="RP23-4M7"
1. 50429
/clone="assembly_fragment"
/clone_end="50429"

```

```

misc_feature
clone_end:50429
vector_side:left

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

```

misc_feature
/clone="assembly_fragment"

```

ORIGIN

Query Match

Best Local Similarity

Matches

18; Conservative

0; Mismatches

0; Indels

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

0; Gaps

```

RESULT 48
AC093466/c 263507 bp DNA linear HTG 23-APR-2003
LOCUS AC093466.2 GI:30018208
DEFINITION Mus musculus clone RP23-4M7, WORKING DRAFT SEQUENCE, 11 unordered
pieces.
AC093466
AC093466.2 GI:30018208
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 263507)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
2 (bases 1 to 263507)
Birren,B., Nuebaum,C. and Lander,E.
Unpublished
REFERENCE
2 (bases 1 to 263507)
Birren,B., Linton,L., Nuebaum,C., Lander,E., Ali,A., Allen,N.,

```

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galsagan, J., Gattaya, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labrecque, K., Lamazet, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Margis, N., Matthews, C., McCarty, M., McEwan, P., McKernan, K., McNetters, R., Meldrum, J., Menus, L., Mihova, T., Mieng, A., Murphy, T., Naylor, J., Nguyen, C., Norou, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, J., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, N., Schupack, R., Seaman, S., Severy, P., Spencer, B., Strang-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talama, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (26-AUG-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 263507)

Anderson, S., Barna, N., Bastien, V., Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galsagan, J., Gattaya, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labrecque, K., Lamazet, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Margis, N., Matthews, C., McCarty, M., McEwan, P., McKernan, K., McNetters, R., Meldrum, J., Menus, L., Mihova, T., Mieng, A., Murphy, T., Naylor, J., Nguyen, C., Norou, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, J., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, N., Schupack, R., Seaman, S., Severy, P., Spencer, B., Strang-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talama, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 17, 2003 this sequence version replaced gi:15290913.

All repeats were identified using RepeatMasker:

Smlt, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W1BR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L13564

Genome Center

Center clone name: 4_M_7

NOTE: This is a 'working draft' sequence. It currently

consists of 11 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 50429: contig of 50429 bp in length

* 50430 50529: gap of 100 bp

* 50530 52201: contig of 1672 bp in length

* 52202 52301: gap of 100 bp

```

* 52302 54425: contig of 2124 bp in length
* 54426 54525: gap of 100 bp
* 54526 57746: contig of 3221 bp in length
* 57747 57846: gap of 100 bp
* 57847 60652: contig of 2806 bp in length
* 60653 60752: gap of 100 bp
* 60753 64812: contig of 4060 bp in length
* 64813 64912: gap of 100 bp
* 64913 79944: contig of 15032 bp in length
* 79945 80044: gap of 100 bp
* 80045 101704: contig of 21660 bp in length
* 101705 101804: gap of 100 bp
* 101805 169526: contig of 67722 bp in length
* 169527 169627: gap of 100 bp
* 169627 203765: contig of 34139 bp in length
* 203766 203865: gap of 100 bp
* 203866 263507: contig of 59642 bp in length.

```

```

FEATURES
source
1..263507
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /clone_id="RP23-4M7"
  /clone_lib="RPC1-23 Female Mouse BAC"
  1..50429
    /note="assembly_fragment"
    clone_end:SP6
    vector_side:left"
    50530..52201
      /note="assembly_fragment"
      52302..54425
        /note="assembly_fragment"
        54526..57746
          /note="assembly_fragment"
          57847..60652
            /note="assembly_fragment"
            60753..64812
              /note="assembly_fragment"
              64913..79944
                /note="assembly_fragment"
                80045..101704
                  /note="assembly_fragment"
                  101805..169526
                    /note="assembly_fragment"
                    169627..203765
                      /note="assembly_fragment"
                      203866..263507
                        /note="assembly_fragment"
                        clone_end:T7
                        vector_side:right"

```

ORIGIN

```

Query Match      90.0%; Score 18; DB 2; Length 263507;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 GGACATGCCCCGGCATGT 18
        |||||
Db      197536 GGACATGCCCCGGCATGT 197519

```

RESULT 49

```

LOCUS      AY395415          1439 bp    DNA          linear    BCT 16-JUN-2004
DEFINITION Unclutered Rubrobacteridae bacterium clone EBI096 16S ribosomal RNA
ACCESSION  AY395415
VERSION     AY395415.1  GI:39546077
KEYWORDS   unclutered Rubrobacteridae bacterium
SOURCE     Bacteria; Actinobacteria; Rubrobacteridae; environmental samples.
REFERENCE  1 (bases 1 to 1439)

```

```

AUTHORS    Schoenborn,L., Salt,M., Hugenholz,P. and Janssen,P.H.
TITLE       Major groups of bacteria in a pasture soil
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1439)
AUTHORS     Schoenborn,L., Salt,M., Hugenholz,P. and Janssen,P.H.
TITLE       Direct Submission
JOURNAL     Submitted (17-SEP-2003) Department of Microbiology and Immunology,
            University of Melbourne, Grattan Street, Parkville, Victoria 3010,
            Australia
FEATURES
source
1..1439
  /organism="unclutered Rubrobacteridae bacterium"
  /mol_type="genomic DNA"
  /isolation_source="pasture soil"
  /db_xref="taxon:211429"
  /clone="EBI096"
  /environmental_sample
  <1..>1439
    /product="16S ribosomal RNA"

```

ORIGIN

```

Query Match      87.0%; Score 17.4; DB 1; Length 1439;
Best Local Similarity 94.7%; Pred. No. 9.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      2 GACATGCCCCGGCATGTC 20
        |||||
Db      969 GACATGACGGCGCATGTC 987

```

```

RESULT 50
LOCUS      AY395415/C          1439 bp    DNA          linear    BCT 16-JUN-2004
DEFINITION Unclutered Rubrobacteridae bacterium clone EBI096 16S ribosomal RNA
ACCESSION  AY395415
VERSION     AY395415
KEYWORDS   unclutered Rubrobacteridae bacterium
SOURCE     unclutered Rubrobacteridae bacterium
ORGANISM   Bacteria; Actinobacteria; Rubrobacteridae; environmental samples.
REFERENCE  1 (bases 1 to 1439)
AUTHORS     Schoenborn,L., Salt,M., Hugenholz,P. and Janssen,P.H.
TITLE       Major groups of bacteria in a pasture soil
JOURNAL     Unpublished
AUTHORS     2 (bases 1 to 1439)
TITLE       Schoenborn,L., Salt,M., Hugenholz,P. and Janssen,P.H.
AUTHORS     Direct Submission
JOURNAL     Submitted (17-SEP-2003) Department of Microbiology and Immunology,
            University of Melbourne, Grattan Street, Parkville, Victoria 3010,
            Australia
FEATURES
source
1..1439
  /organism="unclutered Rubrobacteridae bacterium"
  /mol_type="genomic DNA"
  /isolation_source="pasture soil"
  /db_xref="taxon:211429"
  /clone="EBI096"
  /environmental_sample
  <1..>1439
    /product="16S ribosomal RNA"

```

FEATURES

```

source
1..1439
  /organism="unclutered Rubrobacteridae bacterium"
  /mol_type="genomic DNA"
  /isolation_source="pasture soil"
  /db_xref="taxon:211429"
  /clone="EBI096"
  /environmental_sample
  <1..>1439
    /product="16S ribosomal RNA"

```

ORIGIN

```

Query Match      87.0%; Score 17.4; DB 1; Length 1439;
Best Local Similarity 94.7%; Pred. No. 9.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 GGACATGCCCCGGCATGTC 19
        |||||
Db      987 GACATGACGGCGCATGTC 969

```

Search completed: January 14, 2005, 16:03:44

Job time : 1050.74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 10:09:04 ; Search time 334.737 Seconds
(without alignments)
313.645 Million cell updates/sec

Title: US-09-578-453-2
Perfect score: 20
Sequence: 1 GGACATGCCCGCATGTC 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 500 summaries

Database :

N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAQ88279	AaQ88279 p53 prote
2	20	100.0	20	AAQ88279	AaQ88279 p53 prote
3	20	100.0	20	AAV22590	AaV22590 p53 conse
4	20	100.0	20	AAV22590	AaV22590 p53 conse
5	20	100.0	20	AAV22590	AaV22590 p53 conse
6	20	100.0	20	AAV22590	AaV22590 p53 conse
7	20	100.0	20	AAV22590	AaV22590 p53 conse
8	20	100.0	20	AAV22590	AaV22590 p53 conse
9	20	100.0	20	AAV22590	AaV22590 p53 conse
10	20	100.0	20	AAV22590	AaV22590 p53 conse
11	20	100.0	20	AAV22590	AaV22590 p53 conse
12	20	100.0	20	AAV22590	AaV22590 p53 conse
13	20	100.0	20	AAV22590	AaV22590 p53 conse
14	20	100.0	20	AAV22590	AaV22590 p53 conse
15	20	100.0	20	AAV22590	AaV22590 p53 conse
16	20	100.0	20	AAV22590	AaV22590 p53 conse
17	20	100.0	20	AAV22590	AaV22590 p53 conse
18	20	100.0	20	AAV22590	AaV22590 p53 conse
19	20	100.0	20	AAV22590	AaV22590 p53 conse
20	20	100.0	20	AAV22590	AaV22590 p53 conse
21	20	100.0	20	AAV22590	AaV22590 p53 conse

22	20	100.0	27	AAA61640	AaA61640 p53 bindi
23	20	100.0	56	AAA38055	AaA38055 linker us
24	20	100.0	56	AAA38055	AaA38055 linker us
25	20	100.0	56	ABX16465	AbX16465 selective
26	20	100.0	56	ABX16465	AbX16465 selective
27	20	100.0	56	ABX16450	AbX16450 p53CON-1u
28	20	100.0	56	ABX16450	AbX16450 p53CON-1u
29	20	100.0	56	ABX16449	AbX16449 p53CON-1u
30	20	100.0	56	ABX16449	AbX16449 p53CON-1u
31	20	100.0	56	ABX16449	AbX16449 p53CON-1u
32	20	100.0	56	ABX16466	AbX16466 selective
33	20	100.0	56	ABX16466	AbX16466 selective
34	19	95.0	19	ADBS2736	AdB52736 Oligonuc
35	16.8	84.0	30	AAT32849	AaT32849 Tumour su
36	16.8	84.0	30	AAT32849	AaT32849 Tumour su
37	16.8	84.0	30	AAT32849	AaT32849 Tumour su
38	16.8	84.0	30	AAT32849	AaT32849 Tumour su
39	16.8	84.0	30	AAT32849	AaT32849 Tumour su
40	16.8	84.0	30	AAT32849	AaT32849 Tumour su
41	16.8	84.0	30	AAT32849	AaT32849 Tumour su
42	16.8	84.0	30	AAT32849	AaT32849 Tumour su
43	16.8	84.0	30	AAT32849	AaT32849 Tumour su
44	16.8	84.0	30	AAT32849	AaT32849 Tumour su
45	16.8	84.0	30	AAT32849	AaT32849 Tumour su
46	16.8	84.0	30	AAT32849	AaT32849 Tumour su
47	16.8	84.0	30	AAT32849	AaT32849 Tumour su
48	16.8	84.0	30	AAT32849	AaT32849 Tumour su
49	16.8	84.0	30	AAT32849	AaT32849 Tumour su
50	16.8	84.0	30	AAT32849	AaT32849 Tumour su
51	16.8	84.0	30	AAT32849	AaT32849 Tumour su
52	16.8	84.0	30	AAT32849	AaT32849 Tumour su
53	16.8	84.0	30	AAT32849	AaT32849 Tumour su
54	16.8	84.0	30	AAT32849	AaT32849 Tumour su
55	16.8	84.0	30	AAT32849	AaT32849 Tumour su
56	16.8	84.0	30	AAT32849	AaT32849 Tumour su
57	16.8	84.0	30	AAT32849	AaT32849 Tumour su
58	16.8	84.0	30	AAT32849	AaT32849 Tumour su
59	16.8	84.0	30	AAT32849	AaT32849 Tumour su
60	16.8	84.0	30	AAT32849	AaT32849 Tumour su
61	16.8	84.0	30	AAT32849	AaT32849 Tumour su
62	16.8	84.0	30	AAT32849	AaT32849 Tumour su
63	16.8	84.0	30	AAT32849	AaT32849 Tumour su
64	16.8	84.0	30	AAT32849	AaT32849 Tumour su
65	16.8	84.0	30	AAT32849	AaT32849 Tumour su
66	16.8	84.0	30	AAT32849	AaT32849 Tumour su
67	16.8	84.0	30	AAT32849	AaT32849 Tumour su
68	16.8	84.0	30	AAT32849	AaT32849 Tumour su
69	16.8	84.0	30	AAT32849	AaT32849 Tumour su
70	16.8	84.0	30	AAT32849	AaT32849 Tumour su
71	16.8	84.0	30	AAT32849	AaT32849 Tumour su
72	16.8	84.0	30	AAT32849	AaT32849 Tumour su
73	16.8	84.0	30	AAT32849	AaT32849 Tumour su
74	16.8	84.0	30	AAT32849	AaT32849 Tumour su
75	16.8	84.0	30	AAT32849	AaT32849 Tumour su
76	16.8	84.0	30	AAT32849	AaT32849 Tumour su
77	16.8	84.0	30	AAT32849	AaT32849 Tumour su
78	16.8	84.0	30	AAT32849	AaT32849 Tumour su
79	16.8	84.0	30	AAT32849	AaT32849 Tumour su
80	16.8	84.0	30	AAT32849	AaT32849 Tumour su
81	16.8	84.0	30	AAT32849	AaT32849 Tumour su
82	16.8	84.0	30	AAT32849	AaT32849 Tumour su
83	16.8	84.0	30	AAT32849	AaT32849 Tumour su
84	16.8	84.0	30	AAT32849	AaT32849 Tumour su
85	16.8	84.0	30	AAT32849	AaT32849 Tumour su
86	16.8	84.0	30	AAT32849	AaT32849 Tumour su
87	16.8	84.0	30	AAT32849	AaT32849 Tumour su
88	16.8	84.0	30	AAT32849	AaT32849 Tumour su
89	16.8	84.0	30	AAT32849	AaT32849 Tumour su
90	16.8	84.0	30	AAT32849	AaT32849 Tumour su
91	16.8	84.0	30	AAT32849	AaT32849 Tumour su
92	16.8	84.0	30	AAT32849	AaT32849 Tumour su
93	16.8	84.0	30	AAT32849	AaT32849 Tumour su
94	16.8	84.0	30	AAT32849	AaT32849 Tumour su

95	15.4	77.0	473	6	ABL34893	Abi34893 Murine cD	c 168	15.2	76.0	1920	6	AAD29931	Aad29931 Human cal
96	15.4	77.0	473	6	ABL34893	Abi34893 Murine cD	c 159	15.2	76.0	1924	6	AAS95010	Aas95010 Human DNA
97	15.4	77.0	7282	10	ACC43359	Acc43359 Human NR1	c 170	15.2	76.0	1924	6	AAS95010	Aas95010 Human DNA
98	15.4	77.0	7282	10	ACC43359	Acc43359 Human NR1	c 171	15.2	76.0	1940	9	ADA26336	Ada26336 Human cal
99	15.4	77.0	114411	12	ADQ21090	Adq21090 Human sof	c 172	15.2	76.0	1940	9	ADA26336	Ada26336 Human cal
100	15.4	77.0	114411	12	ADQ21090	Adq21090 Human sof	c 173	15.2	76.0	1958	3	AAA09346	Aaa09346 Human MBP
101	15.2	76.0	20	2	AAVI0231	Aavi0231 p53 bindi	c 174	15.2	76.0	1958	3	AAA09346	Aaa09346 Human MBP
102	15.2	76.0	20	2	AAVI0231	Aavi0231 p53 bindi	c 175	15.2	76.0	1958	6	AAD39469	Aad39469 Human cal
103	15.2	76.0	24	12	ADP26898	Adp26898 Transcrip	c 176	15.2	76.0	1958	6	AAD39469	Aad39469 Human cal
104	15.2	76.0	24	12	ADP26898	Adp26898 Transcrip	c 177	15.2	76.0	1958	6	ADG34537	Adg34537 Human cal
105	15.2	76.0	291	3	AAC18258	Aac18258 Human sec	c 178	15.2	76.0	1958	6	ADG34537	Adg34537 Human cal
106	15.2	76.0	291	3	AAC18258	Aac18258 Human sec	c 179	15.2	76.0	2934	9	ADA48441	Ada48441 Rice gene
107	15.2	76.0	297	6	ABK27635	Abk27635 Human col	c 180	15.2	76.0	2934	9	ADA48441	Ada48441 Rice gene
108	15.2	76.0	297	6	ABK27635	Abk27635 Human col	c 181	15.2	76.0	4067	3	AAA95823	Aaa95823 Human met
109	15.2	76.0	300	6	ABL37865	Abi37865 Human col	c 182	15.2	76.0	4067	3	AAA95823	Aaa95823 Human met
110	15.2	76.0	300	6	ABL37865	Abi37865 Human col	c 183	15.2	76.0	4258	8	ABT18136	Abt18136 Human met
111	15.2	76.0	366	2	AAK27251	Aak27251 Human col	c 184	15.2	76.0	4258	8	ABT18136	Abt18136 Aspergill
112	15.2	76.0	366	2	AAK27251	Aak27251 Human col	c 185	15.2	76.0	4258	8	ABT19950	Abt19950 Aspergill
113	15.2	76.0	401	9	ACH17140	Ach17140 Human adu	c 186	15.2	76.0	4258	8	ABT19950	Abt19950 Aspergill
114	15.2	76.0	401	9	ACH17140	Ach17140 Human adu	c 187	15.2	76.0	5734	10	ADC85382	Adc85382 Human Pcp
115	15.2	76.0	442	2	AAV89143	Aav89143 EST clone	c 188	15.2	76.0	5734	10	ADC85382	Adc85382 Human Pcp
116	15.2	76.0	442	2	AAV89143	Aav89143 EST clone	c 189	15.2	76.0	6942	2	AAQ86480	Aaq86480 Merosin p
117	15.2	76.0	463	10	ADF85659	Adf85659 Human ade	c 190	15.2	76.0	6942	2	AAQ86480	Aaq86480 Merosin p
118	15.2	76.0	463	10	ADF85659	Adf85659 Human ade	c 191	15.2	76.0	9419	3	AAA88894	Aaa88894 Human lam
119	15.2	76.0	472	9	ACH25933	Ach25933 Human adu	c 192	15.2	76.0	9419	3	AAA88894	Aaa88894 Human lam
120	15.2	76.0	472	9	ACH25933	Ach25933 Human adu	c 193	15.2	76.0	9420	3	AAA88892	Aaa88892 Human lam
121	15.2	76.0	491	9	ACH23918	Ach23918 Human adu	c 194	15.2	76.0	9420	3	AAA88892	Aaa88892 Human lam
122	15.2	76.0	491	9	ACH23918	Ach23918 Human adu	c 195	15.2	76.0	9534	2	AAT17419	Aat17419 Merosin c
123	15.2	76.0	492	9	ACH14709	Ach14709 Human adu	c 196	15.2	76.0	9534	2	AAT17419	Aat17419 Merosin c
124	15.2	76.0	492	9	ACH14709	Ach14709 Human adu	c 197	15.2	76.0	9534	2	AAK59768	Aak59768 CDNA enco
125	15.2	76.0	538	12	ACH80089	Ach80089 Human gen	c 198	15.2	76.0	9534	2	AAK59768	Aak59768 CDNA enco
126	15.2	76.0	538	12	ACH80089	Ach80089 Human gen	c 199	15.2	76.0	9534	2	AAK59768	Aak59768 CDNA enco
127	15.2	76.0	569	12	ACH80444	Ach80444 Human gen	c 200	15.2	76.0	9534	3	AAA88893	Aaa88893 Human lam
128	15.2	76.0	569	12	ACH80444	Ach80444 Human gen	c 201	15.2	76.0	9534	3	AAA88893	Aaa88893 Human lam
129	15.2	76.0	574	12	ACH73863	Ach73863 Human gen	c 202	15.2	76.0	9534	6	ABL64110	Abi64110 Breast ca
130	15.2	76.0	574	12	ACH73863	Ach73863 Human gen	c 203	15.2	76.0	9534	6	ABK35565	Abk35565 Gene LAMA
131	15.2	76.0	615	10	ABT32345	Abt32345 Seed deve	c 204	15.2	76.0	9534	6	ABK35565	Abk35565 Gene LAMA
132	15.2	76.0	615	10	ABT32345	Abt32345 Seed deve	c 205	15.2	76.0	9534	6	ABT11110	Abt11110 Human bre
133	15.2	76.0	1024	10	ABZ83272	Abz83272 Toxicolog	c 206	15.2	76.0	9534	6	ABT11110	Abt11110 Human bre
134	15.2	76.0	1024	10	ABZ83272	Abz83272 Toxicolog	c 207	15.2	76.0	9534	10	ADK61038	Adk61038 Ovarian c
135	15.2	76.0	1152	8	ADA70053	Ada70053 Rice gene	c 208	15.2	76.0	9534	10	ADK61038	Adk61038 Ovarian c
136	15.2	76.0	1152	8	ADA70053	Ada70053 Rice gene	c 209	15.2	76.0	9535	3	AAA88891	Aaa88891 Human lam
137	15.2	76.0	1152	11	ADU11565	Adj11565 Rice DNA	c 210	15.2	76.0	9535	3	AAA88891	Aaa88891 Human lam
138	15.2	76.0	1152	11	ADU11565	Adj11565 Rice DNA	c 211	15.2	76.0	9649	12	ADO17880	Ado17880 Human sof
139	15.2	76.0	1251	3	AAA09347	Aaa09347 Human MBP	c 212	15.2	76.0	9649	12	ADO17880	Ado17880 Human sof
140	15.2	76.0	1251	3	AAA09347	Aaa09347 Human MBP	c 213	15.2	76.0	9670	12	ADQ22504	Adq22504 Human sof
141	15.2	76.0	1434	5	AAS91386	Aas91386 DNA enco	c 214	15.2	76.0	9670	12	ADQ22504	Adq22504 Human sof
142	15.2	76.0	1434	5	AAS91386	Aas91386 DNA enco	c 215	15.2	76.0	9746	5	AAS84601	Aas84601 DNA enco
143	15.2	76.0	1436	5	AAS67726	Aas67726 DNA enco	c 216	15.2	76.0	9746	5	AAS84601	Aas84601 DNA enco
144	15.2	76.0	1436	5	AAS67726	Aas67726 DNA enco	c 217	15.2	76.0	10867	4	AAK66345	Aak66345 Human imm
145	15.2	76.0	1481	10	ADG75777	Adg75777 Human pro	c 218	15.2	76.0	10867	4	AAK66345	Aak66345 Human imm
146	15.2	76.0	1481	10	ADG75777	Adg75777 Human pro	c 219	15.2	76.0	10867	4	AAK66345	Aak66345 Human imm
147	15.2	76.0	1890	1	AAN91471	Aan91471 1.9 kb Ro	c 220	15.2	76.0	10867	4	AAK66345	Aak66345 Human imm
148	15.2	76.0	1890	1	AAN91471	Aan91471 1.9 kb Ro	c 221	15.2	76.0	10867	4	AAK66345	Aak66345 Human imm
149	15.2	76.0	1890	10	ADD22456	Ad222456 HLA-B46 T	c 222	15.2	76.0	10867	4	AAK66345	Aak66345 Human imm
150	15.2	76.0	1890	10	ADD22456	Ad222456 HLA-B46 T	c 223	15.2	76.0	10867	8	ABK60593	Abk60593 CDNA enco
151	15.2	76.0	1899	6	ABK11662	Abk11662 DNA enco	c 224	15.2	76.0	10867	8	ABK60593	Abk60593 CDNA enco
152	15.2	76.0	1899	6	ABK11662	Abk11662 DNA enco	c 225	15.2	76.0	10867	10	ADB47307	Adb47307 Human car
153	15.2	76.0	1899	6	ABV73173	Abv73173 Human cal	c 226	15.2	76.0	10867	10	ADB47307	Adb47307 Human car
154	15.2	76.0	1899	10	ADP81442	Adp81442 Leukemia	c 227	15.2	76.0	10867	12	ADJ31343	Adj31343 Human mus
155	15.2	76.0	1899	10	ADP81442	Adp81442 Leukemia	c 228	15.2	76.0	10867	12	ADJ31343	Adj31343 Human mus
156	15.2	76.0	1899	10	ADP81442	Adp81442 Leukemia	c 229	15.2	76.0	10894	4	AAK36611	Aak36611 Human car
157	15.2	76.0	1899	10	ADP81442	Adp81442 Leukemia	c 230	15.2	76.0	10894	4	AAK36611	Aak36611 Human car
158	15.2	76.0	1899	10	ADP81442	Adp81442 Leukemia	c 231	15.2	76.0	10894	4	AAK66343	Aak66343 Human imm
159	15.2	76.0	1899	10	ADP81442	Adp81442 Leukemia	c 232	15.2	76.0	10894	4	AAK66343	Aak66343 Human imm
160	15.2	76.0	1899	10	ADP81442	Adp81442 Leukemia	c 233	15.2	76.0	10894	4	AAK66343	Aak66343 Human imm
161	15.2	76.0	1899	10	ADP81442	Adp81442 Leukemia	c 234	15.2	76.0	10894	10	ADE47305	Ad47305 Human car
162	15.2	76.0	1899	12	ADP62737	Adp62737 Human CDN	c 235	15.2	76.0	10894	10	ADE47305	Ad47305 Human car
163	15.2	76.0	1899	12	ADN03687	Adn03687 Antipsori	c 236	15.2	76.0	10897	4	AAK66344	Aak66344 Human imm
164	15.2	76.0	1899	12	ADN03687	Adn03687 Antipsori	c 237	15.2	76.0	10897	4	AAK66344	Aak66344 Human imm
165	15.2	76.0	1899	12	ADN03687	Adn03687 Antipsori	c 238	15.2	76.0	10897	4	AAK66344	Aak66344 Human imm
166	15.2	76.0	1899	12	ADN03687	Adn03687 Antipsori	c 239	15.2	76.0	10897	4	AAK66344	Aak66344 Human imm
167	15.2	76.0	1920	6	AAD29931	Aad29931 Human cal	c 240	15.2	76.0	10907	4	AAK37603	Aak37603 Human mus

241	15.2	76.0	10907	8	ABX60591	ABX60591 cDNA enco	C 314	14.8	74.0	3934	6	ABK54210	ABK54210 Novel pan
C 242	15.2	76.0	10907	8	ABX60591	ABX60591 cDNA enco	C 315	14.8	74.0	11171	12	ADM80035	ADM80035 Sp1rramyc1
C 243	15.2	76.0	10907	10	ADBE47306	ADBE47306 Human cat	C 316	14.8	74.0	11171	12	ADM80035	ADM80035 Sp1rramyc1
C 244	15.2	76.0	10907	10	ADBE47306	ADBE47306 Human cat	C 317	14.8	74.0	11171	12	ADN97551	ADN97551 S ambofac
C 245	15.2	76.0	10907	12	ADU31341	ADU31341 Human mus	C 318	14.8	74.0	11171	12	ADN97551	ADN97551 S ambofac
C 246	15.2	76.0	10907	12	ADU31341	ADU31341 Human mus	C 319	14.8	74.0	23241	5	AAF97871	AAF97871 Human neu
C 247	15.2	76.0	12718	4	AAK66346	AAK66346 Human imm	C 320	14.8	74.0	23241	5	AAF97871	AAF97871 Human neu
C 248	15.2	76.0	12718	4	AAK66346	AAK66346 Human imm	C 321	14.8	74.0	23241	5	AAF97871	AAF97871 Human neu
C 249	15.2	76.0	12718	4	AAK66346	AAK66346 Human imm	C 322	14.8	74.0	23241	5	AAF97871	AAF97871 Human neu
C 250	15.2	76.0	12718	4	AAK66346	AAK66346 Human imm	C 323	14.8	74.0	42291	9	ACC58253	ACC58253 Clotobioc
C 251	15.2	76.0	12718	8	ABX60595	ABX60595 cDNA enco	C 324	14.8	74.0	42291	9	ACC58253	ACC58253 Clotobioc
C 252	15.2	76.0	12718	8	ABX60595	ABX60595 cDNA enco	C 325	14.8	74.0	63209	4	ABL19214	ABL19214 Drosoph11
C 253	15.2	76.0	12718	12	ADU31345	ADU31345 Human mus	C 326	14.8	74.0	63209	4	ABL19214	ABL19214 Drosoph11
C 254	15.2	76.0	12718	12	ADU31345	ADU31345 Human mus	C 327	14.8	74.0	87350	2	AAK83003	AAK83003 Human WRN
C 255	15.2	76.0	51855	10	ACFP4818	ACFP4818 Melithiaz	C 328	14.8	74.0	87350	2	AAK83003	AAK83003 Human WRN
C 256	15.2	76.0	51855	10	ACFP4818	ACFP4818 Melithiaz	C 329	14.8	74.0	87543	7	ADU84092	ADU84092 Human WRN
C 257	15.2	76.0	67191	9	ADA02903	ADA02903 Mouse Mcm	C 330	14.8	74.0	87543	7	ADU84092	ADU84092 Human WRN
C 258	15.2	76.0	67191	9	ADA02903	ADA02903 Mouse Mcm	C 331	14.8	74.0	187	3	AAF15330	AAF15330 Trichoder
C 259	15.2	76.0	67191	10	ADB72641	ADB72641 Mouse Mcm	C 332	14.4	72.0	187	3	AAF15330	AAF15330 Trichoder
C 260	15.2	76.0	67191	10	ADB72641	ADB72641 Mouse Mcm	C 333	14.4	72.0	370	4	AAH35407	AAH35407 Human col
C 261	15.2	76.0	67191	11	ADL95495	ADL95495 Murine MC	C 334	14.4	72.0	370	4	AAH35407	AAH35407 Human col
C 262	15.2	76.0	67191	11	ADL95495	ADL95495 Murine MC	C 335	14.4	72.0	473	12	ADU42779	ADU42779 Plant CDN
C 263	15.2	76.0	67191	12	ADM74498	ADM74498 Murine ca	C 336	14.4	72.0	473	12	ADU42779	ADU42779 Plant CDN
C 264	15.2	76.0	67191	12	ADM74498	ADM74498 Murine ca	C 337	14.4	72.0	697	12	ADU42781	ADU42781 Plant CDN
C 265	15.2	76.0	110000	4	AAI99682_02	AAI99682_02	C 338	14.4	72.0	697	12	ADU42781	ADU42781 Plant CDN
C 266	15.2	76.0	110000	4	AAI99682_02	AAI99682_02	C 339	14.4	72.0	1150	12	ADU42780	ADU42780 Plant CDN
C 267	15.2	76.0	110000	4	AAI99683_02	AAI99683_02	C 340	14.4	72.0	1150	12	ADU42780	ADU42780 Plant CDN
C 268	15.2	76.0	110000	4	AAI99683_02	AAI99683_02	C 341	14.4	72.0	2002	10	ADG10583	ADG10583 Human STRA
C 269	15.2	76.0	110096	6	ABN95044	ABN95044 Gene #154	C 342	14.4	72.0	2002	10	ADG10583	ADG10583 Human STRA
C 270	15.2	76.0	110096	6	ABN95044	ABN95044 Gene #154	C 343	14.4	72.0	3744	10	ADG10875	ADG10875 Human STRA
C 271	15.2	76.0	173810	6	ABN85752	ABN85752 Mouse chr	C 344	14.4	72.0	3744	10	ADG10875	ADG10875 Human STRA
C 272	15.2	76.0	173810	6	ABN85752	ABN85752 Mouse chr	C 345	14.4	72.0	3901	10	ADG10585	ADG10585 Human STRA
C 273	15.2	76.0	192639	10	ADL13676	ADL13676 Osteoarth	C 346	14.4	72.0	3901	10	ADG10585	ADG10585 Human STRA
C 274	15.2	76.0	192639	10	ADL13676	ADL13676 Osteoarth	C 347	14.4	72.0	4826	10	ADG37607	ADG37607 Human nuc
C 275	15.2	76.0	1327001	12	ADP43517	ADP43517 Human MAD	C 348	14.4	72.0	4826	10	ADG37607	ADG37607 Human nuc
C 276	15.2	76.0	347001	12	ADP43517	ADP43517 Human MAD	C 349	14.4	72.0	5066	10	ADP11506	ADP11506 Murine an
C 277	15.2	76.0	347001	12	ADP43517	ADP43517 Human MAD	C 350	14.4	72.0	5066	10	ADP11506	ADP11506 Murine an
C 278	15.2	76.0	1743	11	ABD14804	ABD14804 Pseudomon	C 351	14.4	72.0	5127	4	AAH18232	AAH18232 Human CDN
C 279	15.2	76.0	1743	11	ABD14804	ABD14804 Pseudomon	C 352	14.4	72.0	5127	4	AAH18232	AAH18232 Human CDN
C 280	15.2	76.0	2199	11	ABD14889	ABD14889 Pseudomon	C 353	14.4	72.0	103599	4	ABX04971	ABX04971 S. cinnam
C 281	15.2	76.0	2199	11	ABD14889	ABD14889 Pseudomon	C 354	14.4	72.0	103599	4	ABX04971	ABX04971 S. cinnam
C 282	15.2	76.0	2433	11	ABD15346	ABD15346 Pseudomon	C 355	14.4	72.0	103599	4	ABX04971	ABX04971 S. cinnam
C 283	15.2	76.0	2433	11	ABD15346	ABD15346 Pseudomon	C 356	14.4	72.0	180385	10	ADL13931	ADL13931 Osteoarth
C 284	14.8	74.0	49	2	AAQ31982	AAQ31982 Synthetic	C 357	14.4	72.0	180385	10	ADL13931	ADL13931 Osteoarth
C 285	14.8	74.0	49	2	AAQ31982	AAQ31982 Synthetic	C 358	14.4	72.0	349980	6	ABQ81842	ABQ81842 Bifidobac
C 286	14.8	74.0	478	9	ACH36044	ACH36044 Human end	C 359	14.4	72.0	349980	6	ABQ81842	ABQ81842 Bifidobac
C 287	14.8	74.0	478	9	ACH36044	ACH36044 Human end	C 360	14.2	71.0	237	4	AAK53818	AAK53818 Murine tr
C 288	14.8	74.0	516	11	ABD06138	ABD06138 Pseudomon	C 361	14.2	71.0	237	4	AAK53818	AAK53818 Murine tr
C 289	14.8	74.0	556	6	ABK45096	ABK45096 cDNA enco	C 362	14.2	71.0	250	6	ABN77384	ABN77384 Human ORF
C 290	14.8	74.0	556	6	ABK45096	ABK45096 cDNA enco	C 363	14.2	71.0	250	6	ABN77384	ABN77384 Human ORF
C 291	14.8	74.0	641	4	AAH56941	AAH56941 P patens	C 364	14.2	71.0	261	10	ADJ34183	ADJ34183 Beta-exon
C 292	14.8	74.0	641	4	AAH56941	AAH56941 P patens	C 365	14.2	71.0	261	10	ADJ34183	ADJ34183 Beta-exon
C 293	14.8	74.0	641	5	AAH50934	AAH50934 Lipid deg	C 366	14.2	71.0	261	10	ADJ34203	ADJ34203 Beta exon
C 294	14.8	74.0	641	5	AAH50934	AAH50934 Lipid deg	C 367	14.2	71.0	261	10	ADJ34203	ADJ34203 Beta exon
C 295	14.8	74.0	747	6	ABK54184	ABK54184 DNA enco	C 368	14.2	71.0	303	11	ABD08441	ABD08441 Pseudomon
C 296	14.8	74.0	747	6	ABK54184	ABK54184 DNA enco	C 369	14.2	71.0	303	11	ABD08441	ABD08441 Pseudomon
C 297	14.8	74.0	755	6	ABN90023	ABN90023 Mouse c10	C 370	14.2	71.0	332	6	ABN23450	ABN23450 Human ORF
C 298	14.8	74.0	755	6	ABN90023	ABN90023 Mouse c10	C 371	14.2	71.0	332	6	ABN23450	ABN23450 Human ORF
C 299	14.8	74.0	762	11	ABD06162	ABD06162 Pseudomon	C 372	14.2	71.0	360	8	ACA31655	ACA31655 Prokaryot
C 300	14.8	74.0	762	11	ABD06162	ABD06162 Pseudomon	C 373	14.2	71.0	360	8	ACA31655	ACA31655 Prokaryot
C 301	14.8	74.0	811	2	AAV47532	AAV47532 cDNA enco	C 374	14.2	71.0	387	6	ABN23646	ABN23646 Human ORF
C 302	14.8	74.0	811	2	AAV47532	AAV47532 cDNA enco	C 375	14.2	71.0	387	6	ABN23646	ABN23646 Human ORF
C 303	14.8	74.0	957	5	AAH88432	AAH88432 DNA enco	C 376	14.2	71.0	393	11	ABD06362	ABD06362 Pseudomon
C 304	14.8	74.0	957	5	AAH88432	AAH88432 DNA enco	C 377	14.2	71.0	393	11	ABD06362	ABD06362 Pseudomon
C 305	14.8	74.0	994	3	AAF07507	AAF07507 Fusaarium	C 378	14.2	71.0	409	4	AAK58992	AAK58992 Human imm
C 306	14.8	74.0	994	3	AAF07507	AAF07507 Fusaarium	C 379	14.2	71.0	409	4	AAK58992	AAK58992 Human imm
C 307	14.8	74.0	1027	4	AAAD04176	AAAD04176 Human Her	C 380	14.2	71.0	456	5	AAH65906	AAH65906 C glucami
C 308	14.8	74.0	1027	4	AAAD04176	AAAD04176 Human Her	C 381	14.2	71.0	456	5	AAH65906	AAH65906 C glucami
C 309	14.8	74.0	1781	5	AAAS00801	AAAS00801 F. venena	C 382	14.2	71.0	530	5	AAH85437	AAH85437 DNA enco
C 310	14.8	74.0	1781	5	AAAS00801	AAAS00801 F. venena	C 383	14.2	71.0	530	5	AAH85437	AAH85437 DNA enco
C 311	14.8	74.0	1949	5	AAAS00797	AAAS00797 F. venena	C 384	14.2	71.0	558	6	ABK53685	ABK53685 Human eos
C 312	14.8	74.0	1949	5	AAAS00797	AAAS00797 F. venena	C 385	14.2	71.0	558	6	ABK53685	ABK53685 Human eos
C 313	14.8	74.0	3934	6	ABK54210	ABK54210 Novel pan	C 386	14.2	71.0	563	12	ADJ42142	ADJ42142 Plant CDN

387	14.2	71.0	566	5	ABA11734	AbA11734 Human ner
388	14.2	71.0	566	3	ABA11734	AbA11734 Human ner
389	14.2	71.0	568	3	AAO1081	AaC01081 Human sec
390	14.2	71.0	568	3	AAO1081	AaC01081 Human sec
391	14.2	71.0	579	4	AAE72100	AaE72100 Corynebacc
392	14.2	71.0	579	4	AAE72100	AaE72100 Corynebacc
393	14.2	71.0	579	4	AAE72106	AaE72106 Corynebacc
394	14.2	71.0	579	4	AAE72106	AaE72106 Corynebacc
395	14.2	71.0	579	4	AAE72092	AaE72092 Corynebacc
396	14.2	71.0	579	4	AAE72092	AaE72092 Corynebacc
397	14.2	71.0	580	12	ADU43658	AdJ43658 Plant CDN
398	14.2	71.0	580	12	ADU43658	AdJ43658 Plant CDN
399	14.2	71.0	589	12	ADJ42130	AdJ42130 Plant CDN
400	14.2	71.0	589	12	ADJ42130	AdJ42130 Plant CDN
401	14.2	71.0	590	5	ABV59325	AbV59325 Human pro
402	14.2	71.0	590	5	ABV59325	AbV59325 Human pro
403	14.2	71.0	591	10	AD122499	Ad122499 Rat liver
404	14.2	71.0	591	10	AD122499	Ad122499 Rat liver
405	14.2	71.0	600	12	ACH69173	Ach69173 Human gen
406	14.2	71.0	600	12	ACH69173	Ach69173 Human gen
407	14.2	71.0	636	3	AAE13652	AaE13652 Aspergill
408	14.2	71.0	636	3	AAE13652	AaE13652 Aspergill
409	14.2	71.0	636	12	ADU42129	AdJ42129 Plant CDN
410	14.2	71.0	636	12	ADU42129	AdJ42129 Plant CDN
411	14.2	71.0	687	8	ACA38025	AcA38025 Pseudomon
412	14.2	71.0	687	8	ACA38025	AcA38025 Pseudomon
413	14.2	71.0	693	11	ABD14590	Abd14590 Pseudomon
414	14.2	71.0	693	11	ABD14590	Abd14590 Pseudomon
415	14.2	71.0	697	6	ABZ78094	AbZ78094 Human bre
416	14.2	71.0	697	6	ABZ78094	AbZ78094 Human bre
417	14.2	71.0	726	11	ABD06393	AbD06393 Pseudomon
418	14.2	71.0	726	11	ABD06393	AbD06393 Pseudomon
419	14.2	71.0	749	4	AAK74242	AaK74242 Human imm
420	14.2	71.0	749	4	AAK74242	AaK74242 Human imm
421	14.2	71.0	749	4	AAK74243	AaK74243 Human imm
422	14.2	71.0	749	4	AAK74243	AaK74243 Human imm
423	14.2	71.0	750	4	AAK74241	AaK74241 Human imm
424	14.2	71.0	750	4	AAK74241	AaK74241 Human imm
425	14.2	71.0	759	11	ABD15110	Abd15110 Pseudomon
426	14.2	71.0	759	11	ABD15110	Abd15110 Pseudomon
427	14.2	71.0	798	11	ABD14825	Abd14825 Pseudomon
428	14.2	71.0	798	11	ABD14825	Abd14825 Pseudomon
429	14.2	71.0	807	6	ABN85857	Abn85857 Nucleotid
430	14.2	71.0	807	6	ABN85857	Abn85857 Nucleotid
431	14.2	71.0	816	11	ABD06347	AbD06347 Pseudomon
432	14.2	71.0	816	11	ABD06347	AbD06347 Pseudomon
433	14.2	71.0	816	11	ABD15310	Abd15310 Pseudomon
434	14.2	71.0	816	11	ABD15310	Abd15310 Pseudomon
435	14.2	71.0	924	8	ACA44058	AcA44058 Pseudomon
436	14.2	71.0	924	8	ACA44058	AcA44058 Pseudomon
437	14.2	71.0	966	11	ABD14833	Abd14833 Pseudomon
438	14.2	71.0	966	11	ABD14833	Abd14833 Pseudomon
439	14.2	71.0	1008	8	ACA45287	AcA45287 Pseudomon
440	14.2	71.0	1008	8	ACA45287	AcA45287 Pseudomon
441	14.2	71.0	1092	6	AB190801	Ab190801 Human pol
442	14.2	71.0	1092	6	AB190801	Ab190801 Human pol
443	14.2	71.0	1093	5	ADM19375	Adm19375 Novel hum
444	14.2	71.0	1093	5	ADM19375	Adm19375 Novel hum
445	14.2	71.0	1125	11	ABD12598	Abd12598 Pseudomon
446	14.2	71.0	1125	11	ABD12598	Abd12598 Pseudomon
447	14.2	71.0	1134	11	ABD14157	Abd14157 Pseudomon
448	14.2	71.0	1134	11	ABD14157	Abd14157 Pseudomon
449	14.2	71.0	1200	6	AAU35969	AaU35969 Corynebacc
450	14.2	71.0	1200	6	AAU35969	AaU35969 Corynebacc
451	14.2	71.0	1254	11	ABD12038	Abd12038 Pseudomon
452	14.2	71.0	1254	11	ABD12038	Abd12038 Pseudomon
453	14.2	71.0	1287	11	AD127124	Ad127124 Human cod
454	14.2	71.0	1287	11	AD127124	Ad127124 Human cod
455	14.2	71.0	1289	3	AACT5770	AaC5770 Human ORF
456	14.2	71.0	1289	3	AACT5770	AaC5770 Human ORF
457	14.2	71.0	1308	11	ABD14010	Abd14010 Pseudomon
458	14.2	71.0	1308	11	ABD14010	Abd14010 Pseudomon
459	14.2	71.0	1326	8	ABT19177	AbT19177 Aspergill

460	14.2	71.0	1326	8	ABT19177	AbT19177 Aspergill
461	14.2	71.0	1326	8	ABT20997	AbT20997 Aspergill
462	14.2	71.0	1326	8	ABT20997	AbT20997 Aspergill
463	14.2	71.0	1344	4	AAE61030	AaE61030 P. putida
464	14.2	71.0	1344	4	AAE61030	AaE61030 P. putida
465	14.2	71.0	1349	2	AAE5764	AaE5764 Mouse non
466	14.2	71.0	1349	2	AAE5764	AaE5764 Mouse non
467	14.2	71.0	1488	11	ABD06334	AbD06334 Pseudomon
468	14.2	71.0	1488	11	ABD06334	AbD06334 Pseudomon
469	14.2	71.0	1611	8	ACA42277	AcA42277 Pseudomon
470	14.2	71.0	1611	8	ACA42277	AcA42277 Pseudomon
471	14.2	71.0	1616	4	AAE6429	AaE6429 Human thr
472	14.2	71.0	1616	4	AAE6429	AaE6429 Human thr
473	14.2	71.0	1668	11	ABD15101	Abd15101 Pseudomon
474	14.2	71.0	1668	11	ABD15101	Abd15101 Pseudomon
475	14.2	71.0	1726	10	ADD47080	AdD47080 Human gen
476	14.2	71.0	1726	10	ADD47080	AdD47080 Human gen
477	14.2	71.0	1740	4	AAE1640	AaE1640 Mycobacte
478	14.2	71.0	1740	4	AAE1640	AaE1640 Mycobacte
479	14.2	71.0	1803	4	AAE3447	AaE3447 Human ADA
480	14.2	71.0	1803	4	AAE3447	AaE3447 Human ADA
481	14.2	71.0	1869	2	AAZ34146	AaZ34146 Human PRO
482	14.2	71.0	1869	2	AAZ34146	AaZ34146 Human PRO
483	14.2	71.0	1869	3	AAE78532	AaE78532 Human PRO
484	14.2	71.0	1869	3	AAE78532	AaE78532 Human PRO
485	14.2	71.0	1869	4	AAE6142	AaE6142 Human DNA
486	14.2	71.0	1869	4	AAE6142	AaE6142 Human DNA
487	14.2	71.0	1869	6	ABK33612	AbK33612 CDNA enco
488	14.2	71.0	1869	6	ABK33612	AbK33612 CDNA enco
489	14.2	71.0	1869	6	ABK33612	AbK33612 CDNA enco
490	14.2	71.0	1869	6	ABK33612	AbK33612 CDNA enco
491	14.2	71.0	1869	6	ABK33612	AbK33612 CDNA enco
492	14.2	71.0	1869	6	ABK33612	AbK33612 CDNA enco
493	14.2	71.0	1869	6	ABK33612	AbK33612 CDNA enco
494	14.2	71.0	1869	8	ACA63714	AcA63714 Novel hum
495	14.2	71.0	1869	8	ACA63714	AcA63714 Novel hum
496	14.2	71.0	1869	8	ACA89592	AcA89592 CDNA enco
497	14.2	71.0	1869	8	ACA89592	AcA89592 CDNA enco
498	14.2	71.0	1869	8	ACA73602	AcA73602 Human sec
499	14.2	71.0	1869	8	ACA73602	AcA73602 Human sec
500	14.2	71.0	1869	8	ACA05917	AcA05917 Human sec

ALIGNMENTS

RESULT 1	AAQ88279 standard; cDNA; 20 BP.
AAQ88279	
XX	AAQ88279;
AC	
XX	
DT	25-MAR-2003 (revised)
DT	06-DEC-1995 (first entry)
XX	
XX	
DE	p53 protein binding site sequence used to treat neurodegeneration.
XX	
XX	
KW	Antisense; recombinant virus; vector: p53; binding site; Huntington's;
KW	neuronal degeneration; Parkinson's; Alzheimer's;
KW	amyotrophic lateral sclerosis; ds.
XX	
OS	Synthetic.
XX	
PN	WO950916-A1.
XX	
PD	13-APR-1995.
XX	
PF	29-SEP-1994; 94WO-FR001142.
XX	
PR	04-OCT-1993; 93FR-00011774.
XX	
PA	(RHON) RHONE POULENC RORER SA.
XX	

PI Mallet J, Revah F, Stutzmann J;
 XX WPI, 1995-155257/20.
 XX
 PT Treating and preventing neurodegenerative diseases - with cpds. that
 PT inhibit p53 activity, partic. an antisense sequence or viral vector.
 XX
 PS Claim 12, Page 11, 21pp; French.
 XX
 CC The sequence of the p53 binding site sequence used in a composition to
 CC treat neuronal degeneration. The sequence may be used as an
 CC oligonucleotide used either solely or as part of a recombinant viral
 CC vector to inhibit the activity of the p53 protein. The oligonucleotide
 CC may be used in conjunction, either as naked DNA or on the same
 CC recombinant viral vector, with a p53 antisense sequence which inhibits
 CC expression of the p53 gene (AA088280). The viral vectors may also encode
 CC a mutated form of p53 able to antagonise p53 protein activity. These
 CC oligonucleotides and recombinant viruses can be used in compositions to
 CC treat and prevent neuronal degeneration, Huntington's, Parkinson's or
 CC Alzheimer's diseases or amyotrophic lateral sclerosis. (Updated on 25-MAR
 CC -2003 to correct PN field.)
 CC
 SQ Sequence 20 BP, 3 A, 7 C, 7 G, 3 T, 0 U, 0 Other;
 Query Match 100.0%; Score 20; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGACATGCCCGGCATGTCC 20
 1 GGACATGCCCGGCATGTCC 20
 DB
 RESULT 2
 AA088279/c
 ID AA088279 standard; cDNA; 20 BP.
 XX
 AC AA088279;
 XX
 DT 25-MAR-2003 (revised)
 DT 06-DEC-1995 (first entry)
 XX
 DE p53 protein binding site sequence used to treat neurodegeneration.
 XX
 KM Antisense; recombinant virus; vector; p53; binding site; Huntington's;
 KM neuronal degeneration; Parkinson's; Alzheimer's;
 KM amyotrophic lateral sclerosis; ds.
 XX
 OS Synthetic.
 XX
 PN WO9509916-A1.
 XX
 PD 13-APR-1995.
 XX
 PF 29-SEP-1994; 94WO-FR001142.
 XX
 PR 04-OCT-1993; 93FR-00011774.
 XX
 PA (RHON) RHONE POULENC RORER SA.
 XX
 PI Mallet J, Revah F, Stutzmann J;
 XX WPI, 1995-155257/20.
 XX
 PT Treating and preventing neurodegenerative diseases - with cpds. that
 PT inhibit p53 activity, partic. an antisense sequence or viral vector.
 XX
 PS Claim 12, Page 11, 21pp; French.
 XX
 CC The sequence of the p53 binding site sequence used in a composition to
 CC treat neuronal degeneration. The sequence may be used as an
 CC oligonucleotide used either solely or as part of a recombinant viral
 CC vector to inhibit the activity of the p53 protein. The oligonucleotide

CC may be used in conjunction, either as naked DNA or on the same
 CC recombinant viral vector, with a p53 antisense sequence which inhibits
 CC expression of the p53 gene (AA088280). The viral vectors may also encode
 CC a mutated form of p53 able to antagonise p53 protein activity. These
 CC oligonucleotides and recombinant viruses can be used in compositions to
 CC treat and prevent neuronal degeneration, Huntington's, Parkinson's or
 CC Alzheimer's diseases or amyotrophic lateral sclerosis. (Updated on 25-MAR
 CC -2003 to correct PN field.)
 CC
 SQ Sequence 20 BP, 3 A, 7 C, 7 G, 3 T, 0 U, 0 Other;
 Query Match 100.0%; Score 20; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGACATGCCCGGCATGTCC 20
 20 GGACATGCCCGGCATGTCC 1
 DB
 RESULT 3
 AA088279/c
 ID AA088279 standard; cDNA; 20 BP.
 XX
 AC AA088279;
 XX
 DT 08-JUL-1998 (first entry)
 DT
 XX
 DE p53 consensus binding sequence.
 XX
 KM R1 subunit; ribonucleotide reductase; cell proliferation; tumour cell;
 KM antisense; growth; inhibition; sensitivity; hydroxyurea; R2 subunit;
 KM chemotherapeutic drug; methotrexate; PAHA; treatment; ss.
 XX
 OS Synthetic.
 XX
 PN WO9805769-A2.
 XX
 PD 12-FEB-1998.
 XX
 PF 01-AUG-1997; 97WO-CA000540.
 XX
 PR 02-AUG-1996; 96US-0023040P.
 PR 07-MAR-1997; 97US-0039595P.
 XX
 PA (GENE-) GENESENSE TECHNOLOGIES INC.
 XX
 PI Wright JA, Young AH;
 XX WPI, 1998-145609/13.
 XX
 PT Antisense oligonucleotides to ribonucleotide reductase genes - used to
 PT modulate tumour growth and inhibit tumour cell proliferation.
 XX
 PS Example 2, Page 24; 79pp; English.
 XX
 CC The present sequence represents a p53 consensus binding sequence. The
 CC specification describes antisense oligonucleotides which are targeted
 CC against the mRNAs of the R1 and R2 subunit sequences of ribonucleotide
 CC reductase. Aberrant expression of the R2 gene, which encodes the second
 CC subunit of the ribonucleotide reductase gene, can determine the malignant
 CC characteristics of cells. Suppression of R2 and R1 gene expression was
 CC found to reduce transformed properties of tumour cells. The antisense
 CC oligonucleotides can be used for modulating tumour cell growth, or for
 CC inhibiting tumour cell proliferation. They can also be used for
 CC increasing the sensitivity of neoplastic cells to chemotherapeutic drugs
 CC (especially to hydroxyurea, methotrexate (MTX), and PAHA). The antisense
 CC oligonucleotides may be used to treat proliferative disorders including
 CC leukaemias, lymphomas, sarcomas, melanomas, various other forms of
 CC cancer, papillomas, arthrosclerosis, psoriasis, polythemia, mastocytosis,
 CC autoimmune diseases, angiogenesis, bacterial infections and viral
 CC infections (including HIV hepatitis, or herpes infections)

SQ Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
1 GGACATGCCCGGCGCATGTCC 20
DB 1 GGACATGCCCGGCGCATGTCC 20

RESULT 4

AAV22590/C
ID AAV22590 standard; DNA; 20 BP.

AC AAV22590;

DT 08-JUL-1998 (first entry)

XX p53 consensus binding sequence.

DE R1 subunit; ribonucleotide reductase; cell proliferation; tumour cell;
XX antisense; growth; inhibition; sensitivity; hydroxyurea; R2 subunit;
KW chemotherapeutic drug; methotrexate; PALA; treatment; ss.

XX Synthetic.

OS WO9805769-A2.

XX 12-FEB-1998.

PD 01-AUG-1997; 97WO-CA000540.

XX 02-AUG-1996; 96US-0023040P.

PR 07-MAR-1997; 97US-003959P.

XX (GENE-) GENESENSE TECHNOLOGIES INC.

XX Wright JA, Young AH;

PI WPI, 1998-145609/13.

XX Antisense oligonucleotides to ribonucleotide reductase genes - used to

PT modulate tumour growth and inhibit tumour cell proliferation.

XX Example 2; Page 24; 79pp; English.

PS The present sequence represents a p53 consensus binding sequence. The

CC specification describes antisense oligonucleotides which are targeted

CC against the mRNAs of the R1 and R2 subunit sequences of ribonucleotide

CC reductase. Aberrant expression of the R2 gene, which encodes the second

CC subunit of the ribonucleotide reductase gene, can determine the malignant

CC characteristics of cells. Suppression of R2 and R1 gene expression was

CC found to reduce transformed properties of tumour cells. The antisense

CC oligonucleotides can be used for modulating tumour cell growth, or for

CC inhibiting tumour cell proliferation. They can also be used for

CC increasing the sensitivity of neoplastic cells to chemotherapeutic drugs

CC (especially to hydroxyurea, methotrexate (MTX), and PALA). The antisense

CC oligonucleotides may be used to treat proliferative disorders including

CC leukaemias, lymphomas, sarcomas, melanomas, various other forms of

CC cancer, papillomas, arteriosclerosis, psoriasis, polythemia, mastocytosis,

CC autoimmune diseases, angiodenesis, bacterial infections and viral

CC infections (including HIV hepatitis, or herpes infections)

CC Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 6.6;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
1 GGACATGCCCGGCGCATGTCC 20
DB 20 GGACATGCCCGGCGCATGTCC 1

RESULT 5
AAK27873
ID AAK27873 standard; DNA; 20 BP.

XX AAK27873;

AC 02-JUN-1999 (first entry)

DT 02-JUN-1999 (first entry)

XX Probe for human CSR protein coding sequence.

DE Cellular stress response protein; CSR1; CSR2; CSR3; human; macrophage;

XX scavenger receptor protein; intracellular stress; arteriosclerosis;

XX diabetic circulatory obstruction; microbial infection; probe; ss.

OS Synthetic.

XX Homo sapiens.

XX WO9909159-A1.

XX 25-FEB-1999.

XX 12-AUG-1998; 98WO-JP003602.

XX 13-AUG-1997; 97JP-00233396.

XX 30-JUL-1998; 98JP-00230121.

XX (NIBS) JAPAN TOBACCO INC.

XX Nakamura Y, Tokino T;

XX WPI, 1999-161032/15.

XX Scavenger receptor proteins - for treatment and diagnosis of disorders

XX involving cell stress.

XX Example 8; Page 159; 175pp; Japanese.

XX This sequence represents a probe for DNA encoding a human cellular stress

XX response (CSR) protein of the invention. The CSR proteins are macrophage

XX scavenger receptor proteins. The CSR proteins can be used in the

XX treatment, gene therapy and diagnosis of diseases in which intracellular

XX stress is important, such as arteriosclerosis, diabetic circulatory

XX obstruction, and microbial infection. Expression of the proteins is

XX induced in vivo in response to intracellular stress, and inhibits cell

XX death as a result of such stress

XX Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 6.6;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
1 GGACATGCCCGGCGCATGTCC 20
DB 1 GGACATGCCCGGCGCATGTCC 20

XX AAK27873/C

XX AAK27873 standard; DNA; 20 BP.

XX AAK27873;

XX 02-JUN-1999 (first entry)

XX Probe for human CSR protein coding sequence.

XX Cellular stress response protein; CSR1; CSR2; CSR3; human; macrophage;

XX scavenger receptor protein; intracellular stress; arteriosclerosis;

XX diabetic circulatory obstruction; microbial infection; probe; ss.

OS Synthetic.
 OS Homo sapiens.
 PN WO9909159-A1.
 XX
 PD 25-FEB-1999.
 XX
 PF 12-AUG-1998; 98WO-JP003602.
 XX
 PR 13-AUG-1997; 97JP-00233396.
 PR 30-JUL-1998; 98JP-00230121.
 XX
 PA (NISR) JAPAN TOBACCO INC.
 XX
 PI Nakamura Y, Tokino T;
 DR WPI; 1999-181032/15.
 XX
 PT Scavenger receptor proteins - for treatment and diagnosis of disorders
 PT involving cell stress.
 XX
 PS Example 8; Page 159; 175pp; Japanese.
 XX
 CC This sequence represents a probe for DNA encoding a human cellular stress
 CC response (CSR) protein of the invention. The CSR proteins are macrophage
 CC scavenger receptor proteins. The CSR proteins can be used in the
 CC treatment, gene therapy and diagnosis of diseases in which intracellular
 CC stress is important, such as arteriosclerosis, diabetic circulatory
 CC obstruction, and microbial infection. Expression of the proteins is
 CC induced in vivo in response to intracellular stress, and inhibits cell
 CC death as a result of such stress
 XX
 SQ Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
 |||||
 DB 20 GGACATGCCCGGCATGTCC 1

RESULT 7
 ID AAA91613 standard; DNA; 20 BP.
 XX
 AC AAA91613;
 XX
 DT 20-DEC-2000 (first entry)
 XX
 DE p53 consensus binding sequence.
 XX
 KM p53 consensus binding sequence; ribonucleotide reductase; R1 protein;
 KM R2 protein; antisense inhibition; cancer; cytostatic; ds.
 XX
 OS Unidentified.
 XX
 PN WO200047733-A1.
 XX
 PD 17-AUG-2000.
 XX
 PF 09-FEB-2000; 2000WO-CA000120.
 XX
 PR 11-FEB-1999; 99US-00249730.
 XX
 PA (GENE-) GENESENSE TECHNOLOGIES INC.
 XX
 PI Wright JA, Young AH;
 XX
 DR WPI; 2000-558216/51.
 XX
 PT New antisense oligonucleotide, AS-I-618-20, is useful for inhibiting

PT tumor cell growth.
 XX
 XX Example 2; Page 71; 137pp; English.
 XX

CC The present sequence is a double-stranded p53 consensus binding sequence.
 CC It was used in an electrophoretic gel mobility shift assay (EMSA) to
 CC determine the presence of wild type p53 in various mouse cell lines. The
 CC sequence was end-labelled with radioactive ATP using T4 polynucleotide
 CC kinase and an excess was incubated with the nuclear lysate from mouse
 CC cells. This was performed as part of a method for modulating cell
 CC proliferation using antisense oligonucleotides directed against the R1
 CC and R2 components of mammalian ribonucleotide reductase. Ribonucleotide
 CC reductase catalyses the conversion of ribonucleotides to their
 CC corresponding deoxyribonucleotides and thus plays an important role in
 CC DNA synthesis and cell proliferation. Regulation of ribonucleotide
 CC reductase is altered in cultured malignant cells and increased levels of
 CC R2 protein and R2 mRNA have been found in pre-malignant and malignant
 CC tissues as compared to normal control tissue samples. The antisense
 CC sequences are therefore useful for inhibiting tumorigenicity of
 CC neoplastic cells and inhibiting metastasis of tumour cells. They are also
 CC useful for increasing sensitivity of neoplastic cells to chemotherapeutic
 CC drugs, thus allowing chemotherapeutic treatments to be used in patients
 CC who have become resistant or less sensitive to chemotherapy
 XX

SQ Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
 |||||
 DB 1 GGACATGCCCGGCATGTCC 20

RESULT 8
 ID AAA91613/c
 ID AAA91613 standard; DNA; 20 BP.

XX
 AC AAA91613;
 XX
 DT 20-DEC-2000 (first entry)
 XX
 DE p53 consensus binding sequence.
 XX
 KM p53 consensus binding sequence; ribonucleotide reductase; R1 protein;
 KM R2 protein; antisense inhibition; cancer; cytostatic; ds.
 XX
 OS Unidentified.
 XX
 PN WO200047733-A1.
 XX
 PD 17-AUG-2000.
 XX
 PF 09-FEB-2000; 2000WO-CA000120.
 XX
 PR 11-FEB-1999; 99US-00249730.
 XX
 PA (GENE-) GENESENSE TECHNOLOGIES INC.
 XX
 PI Wright JA, Young AH;
 XX
 DR WPI; 2000-558216/51.
 XX
 PT New antisense oligonucleotide, AS-I-618-20, is useful for inhibiting
 PT tumor cell growth.
 XX
 PS Example 2; Page 71; 137pp; English.
 XX
 CC The present sequence is a double-stranded p53 consensus binding sequence.
 CC It was used in an electrophoretic gel mobility shift assay (EMSA) to
 CC determine the presence of wild type p53 in various mouse cell lines. The
 CC sequence was end-labelled with radioactive ATP using T4 polynucleotide

kinase and an excess was incubated with the nuclear lysate from mouse cells. This was performed as part of a method for modulating cell proliferation using antisense oligonucleotides directed against the R1 and R2 components of mammalian ribonucleotide reductase. Ribonucleotide reductase catalyzes the conversion of ribonucleotides to their corresponding deoxyribonucleotides and thus plays an important role in DNA synthesis and cell proliferation. Regulation of ribonucleotide reductase is altered in cultured malignant cells and increased levels of R2 protein and R2 mRNA have been found in pre-malignant and malignant tissues as compared to normal control tissue samples. The antisense sequences are therefore useful for inhibiting tumorigenicity of neoplastic cells and inhibiting metastasis of tumour cells. They are also useful for increasing sensitivity of neoplastic cells to chemotherapeutic drugs, thus allowing chemotherapeutic treatments to be used in patients who have become resistant or less sensitive to chemotherapy.

Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 6,6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCATGTCC 20
DB 20 GGACATGCCGGGCATGTCC 1

RESULT 9

AAD27540
ID AAD27540 standard; DNA; 20 BP.

XX AAD27540;

DT 18-APR-2002 (first entry)

DE p53 protein binding consensus sequence, p53CON.

XX p53 protein; pGL3 luciferase reporter vector; luc+; transcription factor;
KM cell cycle control; DNA damage repair; p53CON; apoptosis; ds.

XX Unidentified.

XX WO200196602-A2.

PD 20-DEC-2001.

PF 18-JUN-2001; 2001WO-GB002718.

XX 16-JUN-2000; 2000GB-00014820.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Yang AL, Feasting M;

PI WPI; 2002-130743/17.

PT Determining the p53 status of a sample, useful for assaying for mimetics
or antagonists of p53, or for the presence of DNA damage, comprises
determining whether p53 binds to the pGL3 vector in a sample containing a
pGL3 vector.

PS Disclosure; Page 7; 53pp; English.

XX The patent discloses methods for determining the p53 status of a sample
which comprise providing a sample containing a pGL3 luciferase reporter
vector and determining whether p53 binds to the pGL3 vector. p53 is a
transcription factor that regulates many genes including those associated
with cell cycle control, apoptosis and DNA damage repair. pGL3 reporter
vectors contain a modified firefly luciferase cDNA designated luc+. p53
protein binds to pGL3-basic vector and causes luciferase expression. The
method is useful for determining the p53 status of a sample. It is also
useful for assaying for mimetics or antagonists of p53 and for assaying
for presence of activated p53 and/or DNA damage. The invention also

CC relates to a method of modifying pGL3 vector which involves deletion or
alteration of a CCCGGG motif of the pGL3 vector and/or deleting or
altering a sequence within 20 bp sequence 5' or 3' of CCCGGG motif. The
CC nucleic acid having a sequence incorporating the CCCGGG motif is useful
for conferring promoter activity or p53-responsiveness on a nucleic acid
CC encoding a polypeptide of interest or in assays for p53 transcriptional
CC activity. The present DNA sequence is p53 protein binding consensus
CC sequence containing a CCCGGG motif, p53CON

Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 6,6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCATGTCC 20
DB 1 GGACATGCCGGGCATGTCC 20

RESULT 10

AAD27540/C
ID AAD27540 standard; DNA; 20 BP.

XX AAD27540;

DT 18-APR-2002 (first entry)

DE p53 protein binding consensus sequence, p53CON.

XX p53 protein; pGL3 luciferase reporter vector; luc+; transcription factor;
KM cell cycle control; DNA damage repair; p53CON; apoptosis; ds.

XX Unidentified.

XX WO200196602-A2.

PD 20-DEC-2001.

PF 18-JUN-2001; 2001WO-GB002718.

XX 16-JUN-2000; 2000GB-00014820.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Yang AL, Feasting M;

PI WPI; 2002-130743/17.

PT Determining the p53 status of a sample, useful for assaying for mimetics
or antagonists of p53, or for the presence of DNA damage, comprises
determining whether p53 binds to the pGL3 vector in a sample containing a
pGL3 vector.

PS Disclosure; Page 7; 53pp; English.

XX The patent discloses methods for determining the p53 status of a sample
which comprise providing a sample containing a pGL3 luciferase reporter
vector and determining whether p53 binds to the pGL3 vector. p53 is a
transcription factor that regulates many genes including those associated
with cell cycle control, apoptosis and DNA damage repair. pGL3 reporter
vectors contain a modified firefly luciferase cDNA designated luc+. p53
protein binds to pGL3-basic vector and causes luciferase expression. The
method is useful for determining the p53 status of a sample. It is also
useful for assaying for mimetics or antagonists of p53 and for assaying
for presence of activated p53 and/or DNA damage. The invention also
relates to a method of modifying pGL3 vector which involves deletion or
alteration of a CCCGGG motif of the pGL3 vector and/or deleting or
altering a sequence within 20 bp sequence 5' or 3' of CCCGGG motif. The
nucleic acid having a sequence incorporating the CCCGGG motif is useful
for conferring promoter activity or p53-responsiveness on a nucleic acid
CC encoding a polypeptide of interest or in assays for p53 transcriptional
activity. The present DNA sequence is p53 protein binding consensus

CC sequence containing a CCGGG motif, p53CON
XX
SQ Sequence 20 BP; 3 A; 7 C; 7 G; 3 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGGCATGTCC 20
DB 20 GGACATGCCCGGGCATGTCC 1
RESULT 11
ABK98014
ID ABK98014 standard; DNA; 24 BP.
AC ABK98014;
XX
DT 07-OCT-2002 (first entry)
XX
DE Cell-TRAP method associated Oct oligonucleotide.
XX
KM Transcription factor; transcription factor-responsive element; ds; TFRB;
KM transcription activation; Cell-TRAP.
OS Synthetic.
XX
PN WO200252039-A2.
XX
PD 04-JUL-2002.
XX
PF 21-DEC-2001; 2001WO-CA001861.
XX
PR 27-DEC-2000; 2000CA-02327581.
XX
PA (GENE-) GENEKA BIOTECHNOLOGY INC.
PI Blais Y, Rousseau P, Leblanc B, Camato RN;
XX
DR WPI; 2002-575388/61.
XX
PT A Cell-TRAP method, useful for producing or validating therapeutic
PT compounds, by employing a recombinant cell-based library that carry
PT constructs driven by a minimal promoter and a transcription factor-
PT responsive element.
XX
PS Disclosure; Page 25; 44pp; English.
XX
CC This invention relates to a cell-TRAP method for selecting and producing
CC a therapeutic compound which is presumed selective for, one or a
CC restricted set of given transcriptional pathways and cell types by
CC employing a recombinant cell-based library that carries a construct
CC comprising a reporter gene driven by a minimal promoter and a
CC transcription factor-responsive element (TFRB). The invention also
CC comprises a method for validating a putative compound as a selective
CC therapeutic compound towards a transcription factor response element. The
CC method of the invention is useful for determining the transcriptional
CC activation pathways used by any compound that is biologically active in a
CC cell. This method allows a global view of gene transcription activation
CC in response to diverse stimuli in multiple environments and is a
CC significant improvement over case-by-case approaches, which would be
CC limited to certain aspects of gene activation. It permits to save on
CC clinical trials by screening properly the compounds that would have a
CC lesser probability of providing undesirable, even severe side effects.
CC The present sequence represents a double stranded oligonucleotide probe
CC recognised by a specific transcription factor which is used in the method
CC of the invention
XX
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGGCATGTCC 20
DB 5 GGACATGCCCGGGCATGTCC 24
RESULT 12
ABK98014/C
ID ABK98014 standard; DNA; 24 BP.
AC ABK98014;
XX
DT 07-OCT-2002 (first entry)
XX
DE Cell-TRAP method associated Oct oligonucleotide.
XX
KM Transcription factor; transcription factor-responsive element; ds; TFRB;
KM transcription activation; Cell-TRAP.
OS Synthetic.
XX
PN WO200252039-A2.
XX
PD 04-JUL-2002.
XX
PF 21-DEC-2001; 2001WO-CA001861.
XX
PR 27-DEC-2000; 2000CA-02327581.
XX
PA (GENE-) GENEKA BIOTECHNOLOGY INC.
PI Blais Y, Rousseau P, Leblanc B, Camato RN;
XX
DR WPI; 2002-575388/61.
XX
PT A Cell-TRAP method, useful for producing or validating therapeutic
PT compounds, by employing a recombinant cell-based library that carry
PT constructs driven by a minimal promoter and a transcription factor-
PT responsive element.
XX
PS Disclosure; Page 25; 44pp; English.
XX
CC This invention relates to a cell-TRAP method for selecting and producing
CC a therapeutic compound which is presumed selective for, one or a
CC restricted set of given transcriptional pathways and cell types by
CC employing a recombinant cell-based library that carries a construct
CC comprising a reporter gene driven by a minimal promoter and a
CC transcription factor-responsive element (TFRB). The invention also
CC comprises a method for validating a putative compound as a selective
CC therapeutic compound towards a transcription factor response element. The
CC method of the invention is useful for determining the transcriptional
CC activation pathways used by any compound that is biologically active in a
CC cell. This method allows a global view of gene transcription activation
CC in response to diverse stimuli in multiple environments and is a
CC significant improvement over case-by-case approaches, which would be
CC limited to certain aspects of gene activation. It permits to save on
CC clinical trials by screening properly the compounds that would have a
CC lesser probability of providing undesirable, even severe side effects.
CC The present sequence represents a double stranded oligonucleotide probe
CC recognised by a specific transcription factor which is used in the method
CC of the invention
XX
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGGCATGTCC 20
DB 24 GGACATGCCCGGGCATGTCC 5

RESULT 13
ABK98249
ID ABK98249 standard; DNA; 24 BP.
XX
AC ABK98249,
XX
DT 07-OCT-2002 (first entry)
XX
DE Nucleic acid binding element (NABE) probe #19.
XX
KW Nucleic acid binding element; NABE; nucleic acid binding factor; NABF;
KW probe; ss; nucleic acid binding protein; DNA microarray;
KW viral integration.
XX
OS Synthetic.
XX
PN WO200252037-A2.
XX
PD 04-JUL-2002.
XX
PF 14-DEC-2001; 2001WO-CA001794.
XX
PR 27-DEC-2000; 2000CA-02327561.
XX
PA (GENE-) GENEKA BIOTECHNOLOGY INC.
XX
PI Larose A, Rousseau P, Leblanc B, Camato R;
XX
DR WPI; 2002-575386/61.
XX
PT Screening for nucleic acid binding elements (NABE) or factors (NABF),
PT useful for analyzing DNA-protein interactions, by identifying NABE-NABF
PT complexes using an electromobility shift assay, then identifying NABE
PT using NABE probes.
XX
PS Example 4; Page 22; 37bp; English.
XX
CC The invention relates to a method for screening nucleic acid binding
CC elements (NABEs) that are differentially active in modified cells,
CC comprising identifying complexes between NABEs and nucleic acid binding
CC factors (NABFs) in order to separate and obtain NABEs that bind to NABFs,
CC followed by a hybridisation assay to identify marked NABEs using NABE
CC probes. The method is useful in rapid screening and large-scale
CC identification or analyses of NABEs or NABFs that are differentially
CC active in cells that have undergone particular treatments or that are
CC phenotypically different, using DNA microarray technology. These nucleic
CC acid binding proteins are involved in a variety of cellular processes
CC ranging from transcription and replication to recombination and viral
CC integration. The method is also useful in providing simultaneous analysis
CC of the effect of a given compound on multiple DNA-protein interactions,
CC or for the analysis of compounds that are nucleic acid binding factor
CC analogues, and that selectively bind cis-acting nucleic acids. Sequences
CC ABK98231-ABK98258 represent NABE probes used to make DNA microarrays
XX
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;
XX
QY Query Match 100.0%; Score 20; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 GGACATGCCCGGCGCATGTCC 20
5 GGACATGCCCGGCGCATGTCC 24

XX
DE Nucleic acid binding element (NABE) probe #19.
XX
KW Nucleic acid binding element; NABE; nucleic acid binding factor; NABF;
KW probe; ss; nucleic acid binding protein; DNA microarray;
KW viral integration.
XX
OS Synthetic.
XX
PN WO200252037-A2.
XX
PD 04-JUL-2002.
XX
PF 14-DEC-2001; 2001WO-CA001794.
XX
PR 27-DEC-2000; 2000CA-02327561.
XX
PA (GENE-) GENEKA BIOTECHNOLOGY INC.
XX
PI Larose A, Rousseau P, Leblanc B, Camato R;
XX
DR WPI; 2002-575386/61.
XX
PT Screening for nucleic acid binding elements (NABE) or factors (NABF),
PT useful for analyzing DNA-protein interactions, by identifying NABE-NABF
PT complexes using an electromobility shift assay, then identifying NABE
PT using NABE probes.
XX
PS Example 4; Page 22; 37bp; English.
XX
CC The invention relates to a method for screening nucleic acid binding
CC elements (NABEs) that are differentially active in modified cells,
CC comprising identifying complexes between NABEs and nucleic acid binding
CC factors (NABFs) in order to separate and obtain NABEs that bind to NABFs,
CC followed by a hybridisation assay to identify marked NABEs using NABE
CC probes. The method is useful in rapid screening and large-scale
CC identification or analyses of NABEs or NABFs that are differentially
CC active in cells that have undergone particular treatments or that are
CC phenotypically different, using DNA microarray technology. These nucleic
CC acid binding proteins are involved in a variety of cellular processes
CC ranging from transcription and replication to recombination and viral
CC integration. The method is also useful in providing simultaneous analysis
CC of the effect of a given compound on multiple DNA-protein interactions,
CC or for the analysis of compounds that are nucleic acid binding factor
CC analogues, and that selectively bind cis-acting nucleic acids. Sequences
CC ABK98231-ABK98258 represent NABE probes used to make DNA microarrays
XX
SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;
XX
QY Query Match 100.0%; Score 20; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 GGACATGCCCGGCGCATGTCC 20
24 GGACATGCCCGGCGCATGTCC 5

RESULT 15
ADP26897
ID ADP26897 standard; DNA; 24 BP.
XX
AC ADP26897;
XX
DT 26-AUG-2004 (first entry)
XX
DE Transcription factor activation analysis-related oligonucleotide Segrid1.
XX
KW activation; transcription factor; labelling; ss.
XX
OS Unidentified.
XX
PN JP2004159609-A.

XX 10-JUN-2004.
 PD 15-NOV-2002; 2002JP-00332024.
 XX
 PF 15-NOV-2002; 2002JP-00332024.
 XX
 PR 15-NOV-2002; 2002JP-00332024.
 XX
 PA (FUJI/) FUJIMORI F.
 PA (IKED/) IKEDA H.
 PA (INIT-) INITIUM KK.
 XX
 DR WPI; 2004-424491/40.
 XX
 PT Analyzing grade of activation of transcription factor in test substance,
 PT involves hybridizing probe labeled with specific substance, to test
 PT substance, detecting labeled probe and recovering probe coupled with
 PT transcription factor.
 XX
 PS Example 1; Page 16; 12pp; Japanese.
 XX
 CC This invention relates to a novel method of analyzing the grade of
 CC activation of transcription factor in a test substance, which involves
 CC hybridizing a labelled probe to a test substance, where the probe is
 CC labelled with different labelling substance specific for each test
 CC substance, detecting the labelled probe, and recovering the probe coupled
 CC with transcription factor on a single plate. The invention is useful for
 CC analyzing the grade of activation of transcription factor in a test
 CC substance efficiently and with high sensitivity. The present sequence is
 CC that of an oligonucleotide which was used in the exemplification of the
 CC invention.
 CC
 SO Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;
 XX
 Query Match 100.0%; Score 20; DB 12; Length 24;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGACATGCCCGGCATGTCC 20
 Db 1 GGACATGCCCGGCATGTCC 20
 XX
 RESULT 16
 ADP26897/c
 ID ADP26897 standard; DNA; 24 BP.
 XX
 AC ADP26897;
 XX
 DT 26-AUG-2004 (first entry)
 XX
 DE Transcription factor activation analysis-related oligonucleotide Segid11.
 XX
 KW activation; transcription factor; labelling; ss.
 XX
 OS Unidentified.
 XX
 PN JP2004159609-A.
 XX
 PD 10-JUN-2004.
 XX
 PF 15-NOV-2002; 2002JP-00332024.
 XX
 PR 15-NOV-2002; 2002JP-00332024.
 XX
 PA (FUJI/) FUJIMORI F.
 PA (IKED/) IKEDA H.
 PA (INIT-) INITIUM KK.
 XX
 DR WPI; 2004-424491/40.
 XX
 PT Analyzing grade of activation of transcription factor in test substance,
 PT involves hybridizing probe labeled with specific substance, to test
 PT substance, detecting labeled probe and recovering probe coupled with

PT transcription factor.
 XX
 PS Example 1; Page 16; 12pp; Japanese.
 XX
 CC This invention relates to a novel method of analyzing the grade of
 CC activation of transcription factor in a test substance, which involves
 CC hybridizing a labelled probe to a test substance, where the probe is
 CC labelled with different labelling substance specific for each test
 CC substance, detecting the labelled probe, and recovering the probe coupled
 CC with transcription factor on a single plate. The invention is useful for
 CC analyzing the grade of activation of transcription factor in a test
 CC substance efficiently and with high sensitivity. The present sequence is
 CC that of an oligonucleotide which was used in the exemplification of the
 CC invention.
 CC
 SO Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;
 XX
 Query Match 100.0%; Score 20; DB 12; Length 24;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGACATGCCCGGCATGTCC 20
 Db 20 GGACATGCCCGGCATGTCC 1
 XX
 RESULT 17
 ADP26896
 ID ADP26896 standard; DNA; 24 BP.
 XX
 AC ADP26896;
 XX
 DT 26-AUG-2004 (first entry)
 XX
 DE Transcription factor activation analysis-related oligonucleotide Segid10.
 XX
 KW activation; transcription factor; labelling; ss.
 XX
 OS Unidentified.
 XX
 PN JP2004159609-A.
 XX
 PD 10-JUN-2004.
 XX
 PF 15-NOV-2002; 2002JP-00332024.
 XX
 PR 15-NOV-2002; 2002JP-00332024.
 XX
 PA (FUJI/) FUJIMORI F.
 PA (IKED/) IKEDA H.
 PA (INIT-) INITIUM KK.
 XX
 DR WPI; 2004-424491/40.
 XX
 PT Analyzing grade of activation of transcription factor in test substance,
 PT involves hybridizing probe labeled with specific substance, to test
 PT substance, detecting labeled probe and recovering probe coupled with
 PT transcription factor.
 XX
 PS Example 1; Page 16; 12pp; Japanese.
 XX
 CC This invention relates to a novel method of analyzing the grade of
 CC activation of transcription factor in a test substance, which involves
 CC hybridizing a labelled probe to a test substance, where the probe is
 CC labelled with different labelling substance specific for each test
 CC substance, detecting the labelled probe, and recovering the probe coupled
 CC with transcription factor on a single plate. The invention is useful for
 CC analyzing the grade of activation of transcription factor in a test
 CC substance efficiently and with high sensitivity. The present sequence is
 CC that of an oligonucleotide which was used in the exemplification of the
 CC invention.
 CC
 SO Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 24;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGATGTC 20
 |||||
 Db 5 GGACATGCCCGGCGATGTC 24

RESULT 18
 ADP26896/c
 ID ADP26896 standard; DNA; 24 BP.
 XX
 AC ADP26896;
 XX
 DT 26-AUG-2004 (first entry)
 XX
 DE Transcription factor activation analysis-related oligonucleotide SeqID10.
 XX
 KM activation; transcription factor; labelling; ss.
 XX
 OS Unidentified.
 XX
 PN JP2004159609-A.
 XX
 PD 10-JUN-2004.
 XX
 PF 15-NOV-2002; 2002JP-00332024.
 XX
 PR 15-NOV-2002; 2002JP-00332024.
 XX
 PA (FUJI/) FUJIMORI F.
 PA (IKED/) IKEDA H.
 PA (INIT-) INITIUM KK.
 XX
 DR WPI; 2004-424491/40.
 XX
 PT Analyzing grade of activation of transcription factor in test substance,
 PT involves hybridizing probe labeled with specific substance, to test
 PT substance, detecting labeled probe and recovering probe coupled with
 PT transcription factor.
 XX
 PS Example 1; Page 16; 12pp; Japanese.
 XX
 CC This invention relates to a novel method of analysing the grade of
 CC activation of transcription factor in a test substance, which involves
 CC hybridizing a labelled probe to a test substance, where the probe is
 CC labelled with different labelling substance specific for each test
 CC substance, detecting the labelled probe, and recovering the probe coupled
 CC with transcription factor on a single plate. The invention is useful for
 CC analysing the grade of activation of transcription factor in a test
 CC substance efficiently and with high sensitivity. The present sequence is
 CC that of an oligonucleotide which was used in the exemplification of the
 CC invention.
 CC
 SQ Sequence 24 BP; 4 A; 8 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 24;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGATGTC 20
 |||||
 Db 24 GGACATGCCCGGCGATGTC 5

RESULT 19
 AAA61639
 ID AAA61639 standard; DNA; 26 BP.
 XX
 AC AAA61639;
 XX

DT 23-OCT-2000 (first entry)
 XX
 DE p53 binding site oligonucleotide, SEQ ID NO:40.
 XX
 KM p53 binding site; ubiquitin conjugating enzyme;
 KM ubiquitin mediated proteolysis; human; cellular protein half life;
 KM ubiquitin mediated proteolysis; p53; cyclin; cell cycle regulator;
 KM myc deregulation; human papillomavirus; HPV-18 B6 protein;
 KM cervical cancer; skin cancer; epidermal hyperplasia; epidermal neoplasia;
 KM psoriasis; connective tissue disorder; wound healing; cytostatic;
 KM antiproliferative; anticancer; antipsoriatic; ss.
 XX
 OS Unidentified.
 XX
 PN US6069982-A.
 XX
 PD 30-MAY-2000.
 XX
 PF 17-DEC-1996; 96US-00767942.
 XX
 PR 04-JAN-1994; 94US-00176937.
 PR 23-MAY-1994; 94US-00247904.
 PR 27-MAY-1994; 94US-00250795.
 PR 13-SEP-1994; 94US-00305520.
 PR 07-JUN-1995; 95US-00486663.
 XX
 PA (MITO-) MITOTIX INC.
 XX
 PI Chiu MI, Cottarel G, Berlin V, Draetta G, Damagnez V, Rolfe M;
 XX
 DR WPI; 2000-410854/35.
 XX
 PT Identifying an inhibitor of ubiquitin mediated proteolysis of regulatory
 PT protein for treating cancers involves measuring ubiquitination levels of
 PT the protein in the presence of candidate agent in an eukaryotic cell.
 XX
 PS Example 7; Fig 5; 73pp; English.
 XX
 CC The invention relates to a method of identifying an inhibitor of
 CC ubiquitin-mediated proteolysis of a cell cycle regulatory protein
 CC comprising contacting an engineered eukaryotic cell with a candidate
 CC agent. The eukaryotic cells is engineered to express a recombinant human,
 CC Candida albicans or Schizosaccharomyces pombe ubiquitin-conjugating
 CC enzyme (AAB03169-B03171), a cell cycle regulatory protein (such as p53)
 CC and ubiquitin. The specification also discloses novel Candida albicans
 CC and Schizosaccharomyces pombe ubiquitin-conjugating enzymes, CUBC and
 CC SPUBC (AAB03170, AAB03171), and two novel human ubiquitin-conjugating
 CC enzymes, hUBC and rapUBC (AAB03169, AAB03173). The ubiquitin-mediated
 CC proteolysis system is the major pathway for the selective, controlled
 CC degradation of intracellular proteins in eukaryotic cells. In particular,
 CC this system controls the half-lives of cellular proteins, and is
 CC important in controlling the levels of proteins involved in cell cycle
 CC progression. Alterations in the ubiquitination of these proteins may
 CC therefore play a role in the development of cancers. For example, human
 CC papillomaviruses such as HPV-18 encode a transforming protein, E6
 CC (AAB03176), which combines with a cellular B6-associated protein (B6-AP;
 CC AAB03177) to stimulate the ubiquitination of p53, thus targeting it for
 CC degradation. The ubiquitination inhibitors identified according to the
 CC method of the invention are useful for treatment of cervical cancers and
 CC connective tissue disorders and for controlling the wound healing
 CC process. They are also useful in treatment of hyperplastic epidermal
 CC conditions such as psoriasis, neoplastic epidermal conditions, skin
 CC cancers e.g., basal cell carcinoma, squamous cell carcinoma. The
 CC inhibitors are useful for deregulating myc expression and rendering the
 CC cells sensitive to chemotherapeutic treatment or to upset the balance of
 CC transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-
 CC mediated degradation of cyclins are useful as antiproliferative agents.
 CC Sequences AAA61639-A61640 represent opposite strands of a p53 binding
 CC site oligonucleotide which may be incorporated in a reporter construct
 CC for use in an assay of the invention
 XX
 SQ Sequence 26 BP; 4 A; 10 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 26;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGATGTC 20
 |||||
 DB 6 GGACATGCCCGGCGATGTC 25

RESULT 20
 AAA61639/c
 ID AAA61639 standard; DNA; 26 BP.
 XX
 AC AAA61639;
 XX
 DT 23-OCT-2000 (first entry)
 XX
 DE p53 binding site oligonucleotide, SEQ ID NO:40.
 XX
 KM p53 binding site; ubiquitin conjugating enzyme;
 KM ubiquitin mediated proteolysis; human; cellular protein half life;
 KM ubiquitin mediated proteolysis; p53; cyclin; cell cycle regulator;
 KM myc deregulation; human papillomavirus; HPV-18 B6 protein;
 KM cervical cancer; skin cancer; epidermal hyperplasia; epidermal neoplasia;
 KM psoriasis; connective tissue disorder; wound healing; cytostatic;
 KM antiproliferative; anticancer; antipsoriatic; ss.
 XX
 OS Unidentified.
 XX
 PN US6068982-A.
 XX
 PD 30-MAY-2000.
 XX
 PF 17-DEC-1996; 96US-00767942.
 XX
 PR 04-JAN-1994; 94US-00176937.
 PR 23-MAY-1994; 94US-00247904.
 PR 27-MAY-1994; 94US-00250795.
 PR 13-SEP-1994; 94US-00305520.
 PR 07-JUN-1995; 95US-00486663.
 XX
 PA (MITO-) MITOTIX INC.
 XX
 PI Chiu MI, Cottarel G, Berlin V, Draetta G, Damagnez V, Rolfe M;
 XX
 DR WPI; 2000-410854/35.
 XX
 PT Identifying an inhibitor of ubiquitin mediated proteolysis of regulatory
 PT protein for treating cancers involves measuring ubiquitination levels of
 PT the protein in the presence of candidate agent in an eukaryotic cell.
 XX
 PS Example 7, Fig 5, 73pp; English.
 XX
 CC The invention relates to a method of identifying an inhibitor of
 CC ubiquitin-mediated proteolysis of a cell cycle regulatory protein
 CC comprising contacting an engineered eukaryotic cell with a candidate
 CC agent. The eukaryotic cells is engineered to express a recombinant human,
 CC Candida albicans or Schizosaccharomyces pombe ubiquitin- conjugating
 CC enzyme (AAB03169-B03171), a cell cycle regulatory protein (such as p53)
 CC and ubiquitin. The specification also discloses novel Candida albicans
 CC and Schizosaccharomyces pombe ubiquitin-conjugating enzymes, cdbcs and
 CC subdce (AAB03170, AAB03171), and two novel human ubiquitin-conjugating
 CC enzymes, hdbcs and rapbcs (AAB03169, AAB03173). The ubiquitin-mediated
 CC proteolysis system is the major pathway for the selective, controlled
 CC degradation of intracellular proteins in eukaryotic cells. In particular,
 CC this system controls the half-lives of cellular proteins, and is
 CC important in controlling the levels of proteins involved in cell cycle
 CC progression. Alterations in the ubiquitination of these proteins may
 CC therefore play a role in the development of cancers. For example, human
 CC papillomaviruses such as HPV-18 encode a transforming protein, B6
 CC (AAB03176), which combines with a cellular B6-associated protein (B6-AP;
 CC AAB03177) to stimulate the ubiquitination of p53, thus targeting it for
 CC degradation. The ubiquitination inhibitors identified according to the

CC method of the invention are useful for treatment of cervical cancers and
 CC connective tissue disorders and for controlling the wound healing
 CC process. They are also useful in treatment of hyperplastic epidermal
 CC conditions such as psoriasis, neoplastic epidermal conditions, skin
 CC cancers e.g., basal cell carcinomas, squamous cell carcinomas. The
 CC inhibitors are useful for deregulating myc expression and rendering the
 CC cells sensitive to chemotherapeutic treatment or to upset the balance of
 CC transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-
 CC mediated degradation of cyclins are useful as antiproliferative agents.
 CC Sequences AAA61639-A61640 represent opposite strands of a p53 binding
 CC site oligonucleotide which may be incorporated in a reporter construct
 CC for use in an assay of the invention
 XX
 SQ Sequence 26 BP; 4 A; 10 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 26;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGATGTC 20
 |||||
 DB 25 GGACATGCCCGGCGATGTC 6

RESULT 21
 AAA61640
 ID AAA61640 standard; DNA; 27 BP.
 XX
 AC AAA61640;
 XX
 DT 23-OCT-2000 (first entry)
 XX
 DE p53 binding site oligonucleotide, SEQ ID NO:41.
 XX
 KM p53 binding site; ubiquitin conjugating enzyme;
 KM ubiquitin mediated proteolysis; human; cellular protein half life;
 KM ubiquitin mediated proteolysis; p53; cyclin; cell cycle regulator;
 KM myc deregulation; human papillomavirus; HPV-18 B6 protein;
 KM cervical cancer; skin cancer; epidermal hyperplasia; epidermal neoplasia;
 KM psoriasis; connective tissue disorder; wound healing; cytostatic;
 KM antiproliferative; anticancer; antipsoriatic; ss.
 XX
 OS Unidentified.
 XX
 PN US6068982-A.
 XX
 PD 30-MAY-2000.
 XX
 PF 17-DEC-1996; 96US-00767942.
 XX
 PR 04-JAN-1994; 94US-00176937.
 PR 23-MAY-1994; 94US-00247904.
 PR 27-MAY-1994; 94US-00250795.
 PR 13-SEP-1994; 94US-00305520.
 PR 07-JUN-1995; 95US-00486663.
 XX
 PA (MITO-) MITOTIX INC.
 XX
 PI Chiu MI, Cottarel G, Berlin V, Draetta G, Damagnez V, Rolfe M;
 XX
 DR WPI; 2000-410854/35.
 XX
 PT Identifying an inhibitor of ubiquitin mediated proteolysis of regulatory
 PT protein for treating cancers involves measuring ubiquitination levels of
 PT the protein in the presence of candidate agent in an eukaryotic cell.
 XX
 PS Example 7, Fig 5, 73pp; English.
 XX
 CC The invention relates to a method of identifying an inhibitor of
 CC ubiquitin-mediated proteolysis of a cell cycle regulatory protein
 CC comprising contacting an engineered eukaryotic cell with a candidate
 CC agent. The eukaryotic cells is engineered to express a recombinant human,
 CC Candida albicans or Schizosaccharomyces pombe ubiquitin- conjugating

PA (MITO-) MITOTIX INC.

WO2000022137-A2

PI	Chiu MI, Cottarel G, Berlin V, Draetta G, Damagnez V, Rolfe M,	
XX	WP1; 2000-410854/35.	
XX		
XX	Identifying an inhibitor of ubiquitin mediated proteolysis of regulatory	
PT	protein for treating cancers involves measuring ubiquitination levels of	
PT	the protein in the presence of candidate agent in an eukaryotic cell.	
XX		
PS	Example 7; Fig 5; 73pp; English.	
XX		
PS	The invention relates to a method of identifying an inhibitor of	
CC	ubiquitin-mediated proteolysis of a cell cycle regulatory protein	
CC	comprising contacting an engineered eukaryotic cell with a candidate	
CC	agent. The eukaryotic cells is engineered to express a recombinant human,	
CC	Candida albicans or Schizosaccharomyces pombe ubiquitin-conjugating	
CC	enzyme (AA03169-B03171), a cell cycle regulatory protein (such as p53)	
CC	and ubiquitin. The specification also discloses novel Candida albicans	
CC	and Schizosaccharomyces pombe ubiquitin-conjugating enzymes, caubcs and	
CC	spbcs (AA03170, AA03171), and two novel human ubiquitin-conjugating	
CC	enzyme, hubcs and rapbcs (AA03169, AA03173). The ubiquitin-mediated	
CC	proteolysis system is the major pathway for the selective, controlled	
CC	degradation of intracellular proteins in eukaryotic cells. In particular,	
CC	this system controls the half-lives of cellular proteins, and is	
CC	important in controlling the levels of proteins involved in cell cycle	
CC	progression. Alterations in the ubiquitination of these proteins may	
CC	therefore play a role in the development of cancers. For example, human	
CC	papillomaviruses such as HPV-18 encode a transforming protein, E6	
CC	(AA03176), which combines with a cellular E6-associated protein (E6-AP;	
CC	AA03177) to stimulate the ubiquitination of p53, thus targeting it for	
CC	degradation. The ubiquitination inhibitors identified according to the	
CC	method of the invention are useful for treatment of cervical cancers and	
CC	connective tissue disorders and for controlling the wound healing	
CC	process. They are also useful in treatment of hyperplastic epidermal	
CC	conditions such as psoriasis, neoplastic epidermal conditions, skin	
CC	cancers e.g., basal cell carcinomas, squamous cell carcinomas. The	
CC	inhibitors are useful for deregulating myc expression and rendering the	
CC	cells sensitive to chemotherapeutic treatment or to upset the balance of	
CC	transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-	
CC	mediated degradation of cyclins are useful as antiproliferative agents.	
CC	Sequences AA61839-861640 represent opposite strands of a p53 binding	
CC	site oligonucleotide which may be incorporated in a reporter construct	
CC	for use in an assay of the invention	
XX		
XX	Sequence 27 BP; 4 A; 9 C; 10 G; 4 T; 0 U; 0 Other;	
QY	Query Match	
	Best Local Similarity 100.0%; Score 20; DB 3; Length 27;	
	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	1 GGACATGCCCGGCATGTCC 20	
	26 GGACATGCCCGGCATGTCC 7	
RESULT 23		
AAA38055		
ID	AAA38055 standard; DNA; 56 BP.	
XX		
AC	AAA38055;	
XX		
DT	24-AUG-2000 (first entry)	
XX		
DS	Linker used in p53CON-luciferase plasmid construction.	
XX		
KM	Selective replication; recombinant virus; pathway responsive promoter;	
KM	viral replication repressor; cell death; neoplastic; ablate; TGF-beta;	
KM	transforming growth factor-beta; anti-proliferative; tumour; cancer;	
KM	p53CON; luciferase; ds.	
XX		
OS	Synthetic.	
XX		
XX	MO200022137-A2.	

XX	20-APR-2000.
XX	14-OCT-1999; 99WO-US021452.
XX	15-OCT-1998; 98US-00172686.
XX	(CANU-) CANU1 INC.
XX	Ramachandra M, Shabram PW;
XX	WP1, 2000-317990/27.
XX	
XX	Example 3; Page 37; 49pp; English.
XX	
XX	The present invention relates to a selectively replicating recombinant virus, comprising a pathway responsive promoter, linked to a repressor of viral replication. This sequence represents a linker used in the construction of a luciferase plasmid with p53-responsive promoter p53CON. Replication of the recombinant virus responds to the intracellular conditions of the target cell through the use of a pathway-responsive promoter driving expression of an inhibitor of viral replication. In the target cell, the promoter element is inactive and viral replication is permitted. The recombinant virus is used for killing a target cell in vivo with pathway defect on contact or ex vivo to eliminate tumour cells from stem cell products. The virus is also used for ablating neoplastic cells in a mammalian organism preferably humans, pigs, horses, cattle, dogs or cats in vivo. A recombinant adenoviral vector containing multiple promoters such that the replication of the virus is selectively enabled in cells possessing multiple pathway defects is useful in the treatment of tumours associated with a lack of TGF-beta antiproliferative action. The tumours include breast carcinomas, hepatomas, gastric, colon and skin tumours, as well as B and T lymphomas. The vectors can also be used for the treatment of tumour cells containing p53-pathway defects. They are also useful in treating cancer cells with Rb pathway defects. Alternatively, such vectors may be employed to express a cell surface protein capable of recognition by a binding molecule such as a fluorescently labelled antibody
XX	
XX	Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;
XX	
XX	Query Match 100.0%; Score 20; DB 3; Length 56;
XX	Best Local Similarity 100.0%; Pred. No. 6.7;
XX	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	
XX	1 GGACATGCCCGGCGCATGTCC 20
XX	
XX	7 GGACATGCCCGGCGCATGTCC 26
XX	
XX	RESULT 24
XX	AAA38055/c
XX	ID AAA38055 standard; DNA; 56 BP.
XX	AAA38055;
XX	
XX	24-AUG-2000 (first entry)
XX	
XX	Linker used in p53CON-luciferase plasmid construction.
XX	
XX	Selective replication; recombinant virus, pathway responsive promoter; viral replication repressor; cell death; neoplastic; ablate; TGF-beta; transforming growth factor-beta; anti-proliferative; tumour; cancer; p53CON; luciferase; ds.
XX	
XX	Synthetic.
XX	
XX	WO200022137-A2.
XX	
XX	20-APR-2000.

```

PF 14-OCT-1999; 99WO-US021452.
XX
XX 15-OCT-1998; 98US-00172686.
XX
XX (CANU-) CANU1 INC.
XX
XX Ramachandra M, Shabram PW;
XX WPI; 2000-317990/27.
XX
XX Recombinant viruses which selectively replicate in target cells used in
XX the treatment of tumors comprise a pathway-responsive promoter linked to
XX a viral replication repressor.
XX
XX Example 3; Page 37; 49pp; English.
XX
XX The present invention relates to a selectively replicating recombinant
XX virus, comprising a pathway responsive promoter, linked to a repressor of
XX viral replication. This sequence represents a linker used in the
XX construction of a luciferase plasmid with p53-responsive promoter p53CON.
XX Replication of the recombinant virus responds to the intracellular
XX conditions of the target cell through the use of a pathway-responsive
XX promoter driving expression of an inhibitor of viral replication. In the
XX target cell, the promoter element is inactive and viral replication is
XX permitted. The recombinant virus is used for killing a target cell in
XX vivo with pathway defect on contact or ex vivo to eliminate tumour cells
XX from stem cell products. The virus is also used for ablating neoplastic
XX cells in a mammalian organism preferably humans, pigs, horses, cattle,
XX dogs or cats in vivo. A recombinant adenoviral vector containing multiple
XX promoters such that the replication of the virus is selectively enabled
XX in cells possessing multiple pathway defects is useful in the treatment
XX of tumours associated with a lack of TGF-beta antiproliferative action.
XX The tumours include breast carcinomas, hepatomas, gastric, colon and skin
XX tumours, as well as B and T lymphomas. The vectors can also be used for
XX the treatment of tumour cells containing p53-pathway defects. They are
XX also useful in treating cancer cells with Rb pathway defects.
XX Alternatively, such vectors may be employed to express a cell surface
XX protein capable of recognition by a binding molecule such as a
XX fluorescently labelled antibody
XX
XX Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 3; Length 56;
XX Best Local Similarity 100.0%; Pred. No. 6.7;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGACATGCCCGGGCATGTCC 20
XX |||||||||
XX Db 51 GGACATGCCCGGGCATGTCC 32
XX
XX RESULT 25
XX ID ABX16465 standard; DNA; 56 BP.
XX
XX ABX16465;
XX
XX 15-APR-2003 (first entry)
XX
XX Selectively replicating virus 01/PEME associated PCR primer #5.
XX
XX repressor of viral replication; p53-pathway defect;
XX transforming growth factor beta; TGFbeta; antiproliferative;
XX breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;
XX PCR; primer; ss; 01/PEME.
XX
XX Synthetic.
XX
XX US2002150557-A1.
XX
XX 17-OCT-2002.

```

```

XX 30-JAN-2002; 2002US-00062216.
PF 15-OCT-1998; 98US-0104399P.
PR 13-OCT-1999; 99US-00416812.
XX (RAMA/) RAMACHANDRA M.
PA (SHAB/) SHABRAM P W.
XX Ramachandra M, Shabram PW;
XX WPI; 2003-198263/19.
DR New recombinant virus utilizing a pathway-responsive promoter that
PT inhibits viral replication, useful for diagnosing and treating cancer
PT with p53-pathway defects or with a lack of TGFAproximatelyb
XX antiproliferative action.
XX Example 11; Page 23; 51pp; English.
XX The invention describes a selective replicating recombinant virus (I)
CC comprising a pathway-responsive promoter operably linked to a repressor
CC of viral replication. The methods and compositions are useful for the
CC diagnosis and treatment of disorders containing p53-pathway defects or
CC disorders associated with a lack of transforming growth factor beta
CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon
CC and skin cancer. This sequence represents a PCR primer used in the
CC creation of selectively replicating virus 01/PEME
XX Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 8; Length 56;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGACATGCCCGGCGCATGTCC 20
Db 7 GGACATGCCCGGCGCATGTCC 26
RESULT 26
ABX16465/C
ID ABX16465 standard; DNA; 56 BP.
XX ABX16465;
AC 15-APR-2003 (first entry)
DT Selectively replicating virus 01/PEME associated PCR primer #5.
XX DE Selectively replicating virus 01/PEME associated PCR primer #5.
XX KM Selectively replicating virus; pathway-responsive promoter;
XX KM repressor of viral replication; p53-pathway defect;
XX KM transforming growth factor beta; TGFbeta; antiproliferative;
XX KM breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;
XX KM PCR; primer; ss; 01/PEME.
XX OS Synthetic.
XX US2002150557-A1.
XX 17-OCT-2002.
XX 30-JAN-2002; 2002US-00062216.
XX 15-OCT-1998; 98US-0104399P.
XX PR 13-OCT-1999; 99US-00416812.
XX (RAMA/) RAMACHANDRA M.
XX PA (SHAB/) SHABRAM P W.
XX Ramachandra M, Shabram PW;
XX WPI; 2003-198263/19.

```

```

XX New recombinant virus utilizing a pathway-responsive promoter that
PT inhibits viral replication, useful for diagnosing and treating cancer
PT with p53-pathway defects or with a lack of TGFAproximatelyb
XX antiproliferative action.
XX Example 11; Page 23; 51pp; English.
XX The invention describes a selective replicating recombinant virus (I)
CC comprising a pathway-responsive promoter operably linked to a repressor
CC of viral replication. The methods and compositions are useful for the
CC diagnosis and treatment of disorders containing p53-pathway defects or
CC disorders associated with a lack of transforming growth factor beta
CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon
CC and skin cancer. This sequence represents a PCR primer used in the
CC creation of selectively replicating virus 01/PEME
XX Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 8; Length 56;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGACATGCCCGGCGCATGTCC 20
Db 51 GGACATGCCCGGCGCATGTCC 32
RESULT 27
ABX16450
ID ABX16450 standard; DNA; 56 BP.
XX ABX16450;
AC 15-APR-2003 (first entry)
DT p53CON-luciferase plasmid associated PCR primer #2.
XX DE p53CON-luciferase plasmid associated PCR primer #2.
XX KM Selectively replicating virus; pathway-responsive promoter;
XX KM repressor of viral replication; p53-pathway defect;
XX KM transforming growth factor beta; TGFbeta; antiproliferative;
XX KM breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;
XX KM p53CON-luciferase plasmid; PCR; primer; ss; p53-responsive promoter.
XX OS Synthetic.
XX US2002150557-A1.
XX 17-OCT-2002.
XX 30-JAN-2002; 2002US-00062216.
XX 15-OCT-1998; 98US-0104399P.
XX PR 13-OCT-1999; 99US-00416812.
XX (RAMA/) RAMACHANDRA M.
XX PA (SHAB/) SHABRAM P W.
XX Ramachandra M, Shabram PW;
XX WPI; 2003-198263/19.
XX New recombinant virus utilizing a pathway-responsive promoter that
PT inhibits viral replication, useful for diagnosing and treating cancer
PT with p53-pathway defects or with a lack of TGFAproximatelyb
XX antiproliferative action.
XX Example 3; Page 21; 51pp; English.
XX The invention describes a selective replicating recombinant virus (I)
CC comprising a pathway-responsive promoter operably linked to a repressor
CC of viral replication. The methods and compositions are useful for the
CC diagnosis and treatment of disorders containing p53-pathway defects or

```

CC disorders associated with a lack of transforming growth factor beta
CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon
CC and skin cancer. This sequence represents a PCR primer used in the
CC creation of the p53CON-luciferase plasmid containing a p53-responsive
CC responsive promoter

XX Sequence 56 BP; 10 A; 17 C; 20 G; 9 T; 0 U; 0 Other;

SO Query Match 100.0%; Score 20; DB 8; Length 56;

XX Best Local Similarity 100.0%; Pred. No. 6.7; Mismatches 0; Indels 0; Gaps 0;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGTCC 20

DB 2 GGACATGCCCGGCATGTCC 21

RESULT 28

ABX16450/c

XX ID ABX16450 standard; DNA; 56 BP.

XX AC ABX16450;

XX DT 15-APR-2003 (first entry)

XX DE p53CON-luciferase plasmid associated PCR primer #2.

XX KM Selectively replicating virus; pathway-responsive promoter;

XX KM repressor of viral replication; p53-pathway defect;

XX KM transforming growth factor beta; TGFbeta; antiproliferative;

XX KM breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;

XX KM p53CON-luciferase plasmid; PCR; primer; ss; p53-responsive promoter.

XX OS Synthetic.

XX PN US2002150557-A1.

XX PD 17-OCT-2002.

XX PF 30-JAN-2002; 2002US-00062216.

XX PR 15-OCT-1998; 98US-0104399P.

XX PR 13-OCT-1999; 99US-00416812.

XX PA (RAMA/) RAMACHANDRA M.

XX PA (SHAB/) SHABRAM P W.

XX PI Ramachandra M, Shabram PW;

XX DR WPI; 2003-198263/19.

XX PT New recombinant virus utilizing a pathway-responsive promoter that

XX PT inhibits viral replication, useful for diagnosing and treating cancer

XX PT with p53-pathway defects or with a lack of TGFaproximately

XX PT antiproliferative action.

XX PS Example 3; Page 21; 51pp; English.

XX CC The invention describes a selective replicating recombinant virus (I)

XX CC comprising a pathway-responsive promoter operably linked to a repressor

XX CC of viral replication. The methods and compositions are useful for the

XX CC diagnosis and treatment of disorders containing p53-pathway defects or

XX CC disorders associated with a lack of transforming growth factor beta

XX CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon

XX CC and skin cancer. This sequence represents a PCR primer used in the

XX CC creation of the p53CON-luciferase plasmid containing a p53-responsive

XX CC responsive promoter

XX SO Sequence 56 BP; 10 A; 17 C; 20 G; 9 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 20; DB 8; Length 56;

XX Best Local Similarity 100.0%; Pred. No. 6.7; Mismatches 0; Indels 0; Gaps 0;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGTCC 20

DB 46 GGACATGCCCGGCATGTCC 27

RESULT 29

ABX16449

XX ID ABX16449 standard; DNA; 56 BP.

XX AC ABX16449;

XX DT 15-APR-2003 (first entry)

XX DE p53CON-luciferase plasmid associated PCR primer #1.

XX KM Selectively replicating virus; pathway-responsive promoter;

XX KM repressor of viral replication; p53-pathway defect;

XX KM transforming growth factor beta; TGFbeta; antiproliferative;

XX KM breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;

XX KM p53CON-luciferase plasmid; PCR; primer; ss; p53-responsive promoter.

XX OS Synthetic.

XX PN US2002150557-A1.

XX PD 17-OCT-2002.

XX PF 30-JAN-2002; 2002US-00062216.

XX PR 15-OCT-1998; 98US-0104399P.

XX PR 13-OCT-1999; 99US-00416812.

XX PA (RAMA/) RAMACHANDRA M.

XX PA (SHAB/) SHABRAM P W.

XX PI Ramachandra M, Shabram PW;

XX DR WPI; 2003-198263/19.

XX PT New recombinant virus utilizing a pathway-responsive promoter that

XX PT inhibits viral replication, useful for diagnosing and treating cancer

XX PT with p53-pathway defects or with a lack of TGFaproximately

XX PT antiproliferative action.

XX PS Example 3; Page 21; 51pp; English.

XX CC The invention describes a selective replicating recombinant virus (I)

XX CC comprising a pathway-responsive promoter operably linked to a repressor

XX CC of viral replication. The methods and compositions are useful for the

XX CC diagnosis and treatment of disorders containing p53-pathway defects or

XX CC disorders associated with a lack of transforming growth factor beta

XX CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon

XX CC and skin cancer. This sequence represents a PCR primer used in the

XX CC creation of the p53CON-luciferase plasmid containing a p53-responsive

XX CC responsive promoter

XX SO Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 20; DB 8; Length 56;

XX Best Local Similarity 100.0%; Pred. No. 6.7; Mismatches 0; Indels 0; Gaps 0;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCATGTCC 20

DB 7 GGACATGCCCGGCATGTCC 26

RESULT 30

ABX16449/c

XX ID ABX16449 standard; DNA; 56 BP.

XX AC ABX16449;

```
XX 15-APR-2003 (first entry)
XX
XX p53CON-luciferase plasmid associated PCR primer #1.
DE
XX Selectively replicating virus; pathway-responsive promoter;
XX repressor of viral replication; p53-pathway defect;
XX transforming growth factor beta; TGFbeta; antiproliferative;
XX breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;
XX p53CON-luciferase plasmid; PCR; primer; ss; p53-responsive promoter.
XX
XX Synthetic.
OS
XX US2002150557-A1.
XX
XX 17-OCT-2002.
XX
XX 30-JAN-2002; 2002US-00062216.
XX
XX 15-OCT-1998; 98US-0104399P.
XX 13-OCT-1999; 99US-00416812.
XX
XX (RAMA/) RAMACHANDRA M.
XX (SHAB/) SHABRAM P W.
XX
XX Ramachandra M, Shabram PW;
XX WPI; 2003-198263/19.
XX
XX New recombinant virus utilizing a pathway-responsive promoter that
XX inhibits viral replication, useful for diagnosing and treating cancer
XX with p53-pathway defects or with a lack of TGFbeta; antiproliferative
XX antiproliferative action.
XX
XX Example 3; Page 21; 51pp; English.
XX
XX The invention describes a selective replicating recombinant virus (I)
XX comprising a pathway-responsive promoter operably linked to a repressor
XX of viral replication. The methods and compositions are useful for the
XX diagnosis and treatment of disorders containing p53-pathway defects or
XX disorders associated with a lack of transforming growth factor beta
XX (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon
XX and skin cancer. This sequence represents a PCR primer used in the
XX creation of the p53CON-luciferase plasmid containing a p53-responsive
XX responsive promoter
XX
XX Sequence 56 BP; 9 A; 20 C; 17 G; 10 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 8; Length 56;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGCGCATGTCC 20
DB 51 GGACATGCCCGGCGCATGTCC 32
RESULT 31
ABX16466
ID ABX16466 standard; DNA; 56 BP.
XX
XX ABX16466;
AC
XX
XX 15-APR-2003 (first entry)
XX
XX Selectively replicating virus 01/PEME associated PCR primer #6.
XX
XX Selectively replicating virus; pathway-responsive promoter;
XX repressor of viral replication; p53-pathway defect;
XX transforming growth factor beta; TGFbeta; antiproliferative;
XX breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;
XX PCR; primer; ss; 01/PEME.
XX
```

```
OS Synthetic.
XX
XX US2002150557-A1.
XX
XX 17-OCT-2002.
XX
XX 30-JAN-2002; 2002US-00062216.
XX
XX 15-OCT-1998; 98US-0104399P.
XX 13-OCT-1999; 99US-00416812.
XX
XX (RAMA/) RAMACHANDRA M.
XX (SHAB/) SHABRAM P W.
XX
XX Ramachandra M, Shabram PW;
XX WPI; 2003-198263/19.
XX
XX New recombinant virus utilizing a pathway-responsive promoter that
XX inhibits viral replication, useful for diagnosing and treating cancer
XX with p53-pathway defects or with a lack of TGFbeta; antiproliferative
XX antiproliferative action.
XX
XX Example 11; Page 23-24; 51pp; English.
XX
XX The invention describes a selective replicating recombinant virus (I)
XX comprising a pathway-responsive promoter operably linked to a repressor
XX of viral replication. The methods and compositions are useful for the
XX diagnosis and treatment of disorders containing p53-pathway defects or
XX disorders associated with a lack of transforming growth factor beta
XX (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon
XX and skin cancer. This sequence represents a PCR primer used in the
XX creation of selectively replicating virus 01/PEME
XX
XX Sequence 56 BP; 10 A; 17 C; 19 G; 9 T; 0 U; 1 Other;
SQ
Query Match 100.0%; Score 20; DB 8; Length 56;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGCGCATGTCC 20
DB 2 GGACATGCCCGGCGCATGTCC 21
RESULT 32
ABX16466/c
ID ABX16466 standard; DNA; 56 BP.
XX
XX ABX16466;
AC
XX
XX 15-APR-2003 (first entry)
XX
XX Selectively replicating virus 01/PEME associated PCR primer #6.
XX
XX Selectively replicating virus; pathway-responsive promoter;
XX repressor of viral replication; p53-pathway defect;
XX transforming growth factor beta; TGFbeta; antiproliferative;
XX breast cancer; liver cancer; gastric cancer; colon cancer; skin cancer;
XX PCR; primer; ss; 01/PEME.
XX
XX Synthetic.
OS
XX US2002150557-A1.
XX
XX 17-OCT-2002.
XX
XX 30-JAN-2002; 2002US-00062216.
XX
XX 15-OCT-1998; 98US-0104399P.
XX 13-OCT-1999; 99US-00416812.
XX
XX (RAMA/) RAMACHANDRA M.
XX
```

PA (SHAB/) SHABRAM P W.
 XX Ramachandra M, Shabram PW;
 XX WPI; 2003-198263/19.
 XX
 PT New recombinant virus utilizing a pathway-responsive promoter that
 PT inhibits viral replication, useful for diagnosing and treating cancer
 PT with p53-pathway defects or with a lack of TGFBpropproximately
 PT antiproliferative action.
 XX
 PS Example 11; Page 23-24; 51pp; English.
 XX
 CC The invention describes a selective replicating recombinant virus (I)
 CC comprising a pathway-responsive promoter operably linked to a repressor
 CC of viral replication. The methods and compositions are useful for the
 CC diagnosis and treatment of disorders containing p53-pathway defects or
 CC disorders associated with a lack of transforming growth factor beta
 CC (TGFbeta) antiproliferative action, such as breast, liver, gastric, colon
 CC and skin cancer. This sequence represents a PCR primer used in the
 CC creation of selectively replicating virus 01/PPME
 XX
 SQ Sequence 56 BP; 10 A; 17 C; 19 G; 9 T; 0 U; 1 Other;
 Query Match 100.0%; Score 20; DB 8; Length 56;
 Best Local Similarity 100.0%; Pred. No. 6.7;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGCCCCGGCATGTCC 20
 Db 46 GACATGCCCCGGCATGTCC 27
 XX
 RESULT 33
 ADE52736
 ID ADE52736 standard; DNA; 19 BP.
 XX
 AC ADE52736;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Oligonucleotide SEQ ID 102.
 XX
 KM DNA-binding protein; interferon-activatable protein; ss.
 XX
 OS Synthetic.
 XX
 PN WO2003089466-A1.
 PD 30-OCT-2003.
 XX
 PF 18-APR-2003; 2003WO-JP004981.
 XX
 PR 19-APR-2002; 2002JP-00117840.
 PR 30-APR-2002; 2002JP-00128418.
 PR 30-APR-2002; 2002JP-00128779.
 PR 04-DEC-2002; 2002JP-00352469.
 XX
 PA (RIKE) RIKEN KK.
 PA (DNAF-) DNAFORM KK.
 PA (MITU) MITSUBISHI CHEM CORP.
 XX
 PI Hayaishizaki Y, Kamiya M, Kubodera H;
 XX WPI; 2004-011681/01.
 DR
 XX
 PT Proteins with DNA binding activity and substances that affect their
 PT activity or expression, useful for treating associated disorders.
 XX
 PS Example 9; SEQ ID NO 102; 237pp; Japanese.
 XX
 CC The present invention relates to novel proteins (ADE52648-ADE52660,
 CC ADE52670 and ADE52672) and their coding sequences (ADE52635-ADE52647,
 CC ADE52670 and ADE52672) and their coding sequences (ADE52635-ADE52647,

CC ADE52669 and ADE52671). The proteins have a DNA-binding activity or an
 CC interferon-activatable protein (IAP)-like activity. The present
 CC oligonucleotide is related to p53.
 XX
 SQ Sequence 19 BP; 3 A; 6 C; 7 G; 3 T; 0 U; 0 Other;
 Query Match 95.0%; Score 19; DB 12; Length 19;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGCCCCGGCATGTTC 19
 Db 1 GACATGCCCCGGCATGTTC 19
 XX
 RESULT 34
 ADE52736/C
 ID ADE52736 standard; DNA; 19 BP.
 XX
 AC ADE52736;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Oligonucleotide SEQ ID 102.
 XX
 KM DNA-binding protein; interferon-activatable protein; ss.
 XX
 OS Synthetic.
 XX
 PN WO2003089466-A1.
 PD 30-OCT-2003.
 XX
 PF 18-APR-2003; 2003WO-JP004981.
 XX
 PR 19-APR-2002; 2002JP-00117840.
 PR 30-APR-2002; 2002JP-00128418.
 PR 30-APR-2002; 2002JP-00128779.
 PR 04-DEC-2002; 2002JP-00352469.
 XX
 PA (RIKE) RIKEN KK.
 PA (DNAF-) DNAFORM KK.
 PA (MITU) MITSUBISHI CHEM CORP.
 XX
 PI Hayaishizaki Y, Kamiya M, Kubodera H;
 XX WPI; 2004-011681/01.
 DR
 XX
 PT Proteins with DNA binding activity and substances that affect their
 PT activity or expression, useful for treating associated disorders.
 XX
 PS Example 9; SEQ ID NO 102; 237pp; Japanese.
 XX
 CC The present invention relates to novel proteins (ADE52648-ADE52660,
 CC ADE52670 and ADE52672) and their coding sequences (ADE52635-ADE52647,
 CC ADE52669 and ADE52671). The proteins have a DNA-binding activity or an
 CC interferon-activatable protein (IAP)-like activity. The present
 CC oligonucleotide is related to p53.
 XX
 SQ Sequence 19 BP; 3 A; 6 C; 7 G; 3 T; 0 U; 0 Other;
 Query Match 95.0%; Score 19; DB 12; Length 19;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GACATGCCCCGGCATGTTC 20
 Db 19 GACATGCCCCGGCATGTTC 1
 XX
 RESULT 35
 AAT32849
 ID AAT32849 standard; DNA; 30 BP.

```

XX AC AAT32849;
XX XX
XX DT 06-NOV-1996 (first entry)
XX XX
XX DE Tumour suppressor p53 binding site oligonucleotide BC.
XX XX
XX KM p53 protein; tumour suppressor; tetramerisation domain;
XX KM chimaeric protein; GCN4; transcription factor; leucine zipper;
XX KM gene therapy; vector; cell proliferation; cancer; apoptosis;
XX KM autoimmune disease; immune tolerance; ds.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT repeat_unit 3..7
XX FT /tag= a
XX FT repeat_unit 8..12
XX FT /tag= b
XX FT repeat_unit 13..17
XX FT /tag= c
XX FT repeat_unit 18..22
XX FT /tag= d
XX XX
XX PM MO9616989-A1.
XX PD 06-JUN-1996.
XX XX
XX PF 27-NOV-1995; 95WO-US015353.
XX XX
XX PR 28-NOV-1994; 94US-00347792.
XX PR 28-APR-1995; 95US-00431357.
XX PR 01-JUN-1995; 95US-00456623.
XX XX
XX PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX XX
XX PI Halazonecis TD;
XX XX
XX DR WPI; 1996-286828/29.
XX XX
XX PT New chimaeric p53 protein with heterologous tetramerisation domain - and
XX PT related DNA and vectors, useful for treating abnormal cell proliferation,
XX PT esp. cancer, auto-immune disease, etc.
XX PS Example 1; Page 45; 123pp; English.
XX XX
XX CC Oligonucleotides BC.V4A (AAT32848) and BC (AAT32849) contain artificial
XX CC pentanucleotide repeat sites recognised by human tumour suppressor
XX CC protein p53 (see also AAW02617). They were used with oligonucleotide
XX CC Ewafi (AAT32847), which corresponds to the enhancer that drives p53-
XX CC dependent transcription of the waf1 gene, to assay p53-mediated
XX CC transcriptional activity and the DNA binding activity of wild-type p53
XX CC and of p53 proteins with altered tetramerisation domains
XX XX
XX SQ Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
XX
Query Match 84.0%; Score 16.8; DB 2; Length 30;
Best Local Similarity 90.0%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGCGCATGTC 20
Db 3 GGGCATGTCCGGGCGCATGTCC 22

```

```

DE DE Tumour suppressor p53 binding site oligonucleotide BC.
XX XX
XX KM p53 protein; tumour suppressor; tetramerisation domain;
XX KM chimaeric protein; GCN4; transcription factor; leucine zipper;
XX KM gene therapy; vector; cell proliferation; cancer; apoptosis;
XX KM autoimmune disease; immune tolerance; ds.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT repeat_unit 3..7
XX FT /tag= a
XX FT repeat_unit 8..12
XX FT /tag= b
XX FT repeat_unit 13..17
XX FT /tag= c
XX FT repeat_unit 18..22
XX FT /tag= d
XX XX
XX PM MO9616989-A1.
XX PD 06-JUN-1996.
XX XX
XX PF 27-NOV-1995; 95WO-US015353.
XX XX
XX PR 28-NOV-1994; 94US-00347792.
XX PR 28-APR-1995; 95US-00431357.
XX PR 01-JUN-1995; 95US-00456623.
XX XX
XX PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX XX
XX PI Halazonecis TD;
XX XX
XX DR WPI; 1996-286828/29.
XX XX
XX PT New chimaeric p53 protein with heterologous tetramerisation domain - and
XX PT related DNA and vectors, useful for treating abnormal cell proliferation,
XX PT esp. cancer, auto-immune disease, etc.
XX PS Example 1; Page 45; 123pp; English.
XX XX
XX CC Oligonucleotides BC.V4A (AAT32848) and BC (AAT32849) contain artificial
XX CC pentanucleotide repeat sites recognised by human tumour suppressor
XX CC protein p53 (see also AAW02617). They were used with oligonucleotide
XX CC Ewafi (AAT32847), which corresponds to the enhancer that drives p53-
XX CC dependent transcription of the waf1 gene, to assay p53-mediated
XX CC transcriptional activity and the DNA binding activity of wild-type p53
XX CC and of p53 proteins with altered tetramerisation domains
XX XX
XX SQ Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
XX
Query Match 84.0%; Score 16.8; DB 2; Length 30;
Best Local Similarity 90.0%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGCGCATGTC 20
Db 22 GGACATGCCCGGCGCATGTC 3

```

```

RESULT 37
AAT39778
ID AAT39778 standard; DNA; 30 BP.
XX
XX AC AAT39778;
XX XX
XX DT 30-APR-1997 (first entry)
XX XX
XX DE Competitor DNA oligonucleotide BC.
XX XX
XX KM Human; p53; cell proliferation; cell death; regulator; tumour; psoriasis;
XX KM negative regulatory region; DNA damaging agent; transplant rejection;
XX KM abnormal cell proliferation; atherosclerosis; cancer; autoimmune disease;

```


KW arterial restenosis; immune response; apoptosis; inducer; therapy;
KW proliferating lymphocytes; ds.
XX
OS Homo sapiens.
XX
PN MO9625434-A1.
XX
PD 22-AUG-1996.
XX
PF 16-FEB-1996; 96WO-US001535.
XX
PR 16-FEB-1995; 95US-00392542.
XX
PA (WIST-) WISTAR INST.
XX (FARB) BAYER CORP.
XX
PI Halazonetis T, Hartwig W;
XX
DR WPI, 1996-393345/39.
XX
PT New human p53-isomorph peptide(s) and peptide:mimetic cpds. - used for
PT activating p53 function, e.g. for treating tumours, cancers, psoriasis,
PT etc.
XX
PS Example 2; Page 24; 55pp; English.
XX
CC AA139774-139778 represent competitor DNA's for the peptides of the
CC invention (see AAM05350-W05374). The peptides of the invention are p53
CC fragments (see AAM05344 for full length wild type sequence). The p53
CC protein functions to regulate cell proliferation and cell death, and is
CC mutated in more than half of all human tumours. The peptide sequences are
CC used to activate the DNA binding activity of wild type p53, and p53
CC mutants (see AAM05345-W05349). The peptides of the invention consist of
CC at least four sequential amino acids from a negative regulatory region
CC which maps to residues 361-383 of p53. The peptide sequences preferably
CC contain four amino acids from a non-human p53 sequence, contain D-form
CC amino acids, and can also be cyclic peptides. The sequences retain the
CC structural characteristics of the original peptides, but the
CC modifications render them less susceptible to cleavage by proteases and
CC exopeptidases. As these sequences activate p53 DNA binding, they can be
CC used to identify p53 mutants. The peptides can also be used for treating
CC a patient with a tumour expressing a p53 mutant whose ability to bind DNA
CC may be activated by one of the peptides. They can also be used for
CC treating conditions such as exposure to DNA damaging agents, abnormal
CC cell proliferation characteristic of psoriasis, atherosclerosis, cancer,
CC arterial restenosis, autoimmune diseases and undesirable immune responses
CC accompanying rejection of a transplant. The peptides can also induce
CC apoptosis of specific cells, such as proliferating lymphocytes
XX
SQ Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 2; Length 30;
Best Local Similarity 90.0%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGCGCATGCC 20
Db 3 GGCGATGTCGCGCGCATGCC 22
RESULT 38
AA139778/c
ID AA139778 standard; DNA; 30 BP.
XX
AC AA139778;
XX
DT 30-APR-1997 (first entry)
XX
DB Competitor DNA oligonucleotide BC.
XX
KW Human; p53; cell proliferation; cell death; regulator; tumour; psoriasis;
KW negative regulatory region; DNA damaging agent; transplant rejection;
KW abnormal cell proliferation; atherosclerosis; cancer; autoimmune disease;

KW arterial restenosis; immune response; apoptosis; inducer; therapy;
KW proliferating lymphocytes; ds.
XX
OS Homo sapiens.
XX
PN MO9625434-A1.
XX
PD 22-AUG-1996.
XX
PF 16-FEB-1996; 96WO-US001535.
XX
PR 16-FEB-1995; 95US-00392542.
XX
PA (WIST-) WISTAR INST.
XX (FARB) BAYER CORP.
XX
PI Halazonetis T, Hartwig W;
XX
DR WPI, 1996-393345/39.
XX
PT New human p53-isomorph peptide(s) and peptide:mimetic cpds. - used for
PT activating p53 function, e.g. for treating tumours, cancers, psoriasis,
PT etc.
XX
PS Example 2; Page 24; 55pp; English.
XX
CC AA139774-139778 represent competitor DNA's for the peptides of the
CC invention (see AAM05350-W05374). The peptides of the invention are p53
CC fragments (see AAM05344 for full length wild type sequence). The p53
CC protein functions to regulate cell proliferation and cell death, and is
CC mutated in more than half of all human tumours. The peptide sequences are
CC used to activate the DNA binding activity of wild type p53, and p53
CC mutants (see AAM05345-W05349). The peptides of the invention consist of
CC at least four sequential amino acids from a negative regulatory region
CC which maps to residues 361-383 of p53. The peptide sequences preferably
CC contain four amino acids from a non-human p53 sequence, contain D-form
CC amino acids, and can also be cyclic peptides. The sequences retain the
CC structural characteristics of the original peptides, but the
CC modifications render them less susceptible to cleavage by proteases and
CC exopeptidases. As these sequences activate p53 DNA binding, they can be
CC used to identify p53 mutants. The peptides can also be used for treating
CC a patient with a tumour expressing a p53 mutant whose ability to bind DNA
CC may be activated by one of the peptides. They can also be used for
CC treating conditions such as exposure to DNA damaging agents, abnormal
CC cell proliferation characteristic of psoriasis, atherosclerosis, cancer,
CC arterial restenosis, autoimmune diseases and undesirable immune responses
CC accompanying rejection of a transplant. The peptides can also induce
CC apoptosis of specific cells, such as proliferating lymphocytes
XX
SQ Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 2; Length 30;
Best Local Similarity 90.0%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGCGCATGCC 20
Db 22 GGCGATGTCGCGCGCATGCC 3
RESULT 39
AA160143
ID AA160143 standard; DNA; 30 BP.
XX
AC AA160143;
XX
DT 25-JUN-1997 (first entry)
XX
DB Oligonucleotide BC for testing DNA binding activity.
XX
KW p53; tumour suppressor; cancer; therapy; cell proliferation; apoptosis;
KW protein engineering; DNA binding; ds.

```

OS Synthetic.
XX
XX Key Location/Qualifiers
XX misc_binding 3..22
XX /tag= a
XX /note= "p53 DNA binding site"
XX
XX WO9710843-A1.
XX
XX
XX 27-MAR-1997.
XX
XX 20-SEP-1996; 96WO-US015188.
XX
XX 22-SEP-1995; 95US-0004802P.
XX
XX 21-AUG-1996; 96US-00697221.
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX Halazonetis TD;
XX
XX WPI; 1997-202618/18.
XX
XX R284K modified p53 protein having DNA binding ability - useful in
XX treatment of cancer.
XX
XX Example 2; Page 20; 82pp; English.
XX
XX Oligonucleotide BC (AAT60143) contains a high affinity p53 binding site,
XX and was used with oligonucleotides Bp21 (AAT60144), Bgadd45 (AAT60145)
XX and non-specific TF3 (AAT60146) to examine the DNA binding activity of
XX p53 variants carrying a Thr284 to Arg subseq. The results showed that the
XX T284R subseq. enhanced binding to all p53 proteins examined. For wild-
XX type p53 (see also AAW13949), the effect was evident with
XX oligonucleotides BC and Bp21, for p53Gln248 (see also AAW13968) it was
XX evident for oligonucleotide BC, and for p53His273 (see also AAW13969) and
XX p53Cys273 (see also AAW13970) it was evident for oligonucleotides BC,
XX Bp21 and Bgadd45
XX
XX Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
XX
SQ
XX
XX Query Match 84.0%; Score 16.8; DB 2; Length 30;
XX Best Local Similarity 90.0%; Pred. No. 2.2e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 GGACATGCCCGGCGCATGTC 20
XX ||||| ||||| |||||
XX 3 GGCGATGTCGCGGCGCATGTC 22
XX
XX RESULT 40
XX AAT60143/c
XX ID AAT60143 standard; DNA; 30 BP.
XX
XX AAT60143;
XX
XX 25-JUN-1997 (first entry)
XX
XX Oligonucleotide BC for testing DNA binding activity.
XX
XX p53; tumour suppressor; cancer; therapy; cell proliferation; apoptosis;
XX protein engineering; DNA binding; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX misc_binding 3..22
XX /tag= a
XX /note= "p53 DNA binding site"
XX
XX WO9710843-A1.
XX
XX 27-MAR-1997.
XX

```

```

PF 20-SEP-1996; 96WO-US015188.
XX
XX 22-SEP-1995; 95US-0004802P.
XX
XX 21-AUG-1996; 96US-00697221.
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX Halazonetis TD;
XX
XX WPI; 1997-202618/18.
XX
XX R284K modified p53 protein having DNA binding ability - useful in
XX treatment of cancer.
XX
XX Example 2; Page 20; 82pp; English.
XX
XX Oligonucleotide BC (AAT60143) contains a high affinity p53 binding site,
XX and was used with oligonucleotides Bp21 (AAT60144), Bgadd45 (AAT60145)
XX and non-specific TF3 (AAT60146) to examine the DNA binding activity of
XX p53 variants carrying a Thr284 to Arg subseq. The results showed that the
XX T284R subseq. enhanced binding to all p53 proteins examined. For wild-
XX type p53 (see also AAW13949), the effect was evident with
XX oligonucleotides BC and Bp21, for p53Gln248 (see also AAW13968) it was
XX evident for oligonucleotide BC, and for p53His273 (see also AAW13969) and
XX p53Cys273 (see also AAW13970) it was evident for oligonucleotides BC,
XX Bp21 and Bgadd45
XX
XX Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 U; 0 Other;
XX
SQ
XX
XX Query Match 84.0%; Score 16.8; DB 2; Length 30;
XX Best Local Similarity 90.0%; Pred. No. 2.2e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 GGACATGCCCGGCGCATGTC 20
XX ||||| ||||| |||||
XX 22 GGACATGCCCGGCGCATGTC 3
XX
XX RESULT 41
XX ADU60926
XX ID ADU60926 standard; DNA; 4157 BP.
XX
XX ADU60926;
XX
XX 06-MAY-2004 (first entry)
XX
XX Concatemer of PDB4A oligonucleotides.
XX
XX interleukin; IL-4 receptor; IL-5 receptor; lung disease;
XX airway inflammation; allergy; asthma; impeded respiration;
XX cystic fibrosis; acute respiratory distress syndrome;
XX pulmonary hypertension; lung inflammation; bronchitis; ds.
XX
XX Homo sapiens.
XX
XX WO2004011613-A2.
XX
XX 05-FEB-2004.
XX
XX 25-JUL-2003; 2003WO-US023509.
XX
XX 29-JUL-2002; 2002US-0399076P.
XX
XX (EPIC-) EPICGENESIS PHARM INC.
XX
XX Nyce JW, Tang L, Sandrasagra A, Aguilar D, Miller S;
XX Shahabuddin S, Lu H, Cong H;
XX WPI; 2004-203534/19.
XX
XX Novel single or multiple target oligonucleotide anti-sense to e.g.
XX initiation codons and introns of respiratory disease-relevant genes e.g.,
XX CCR1, RANTES, MCP4, useful for prophylaxis or treating respiratory
XX

```

PT disease e.g., asthma.
XX
PS Claim 2; SEQ ID NO 1782; 85pp; English.
XX
CC The present invention relates to an oligonucleotide anti-sense to e.g.,
CC initiation codon, coding region with 2-10 nucleotides of 5'-end and 3'-
CC end of nucleic acid target comprising gene(s) chosen from e.g.,
CC interleukin (IL)-4 receptor, IL-5 receptor or salts of the
CC oligonucleotide and optionally surfactant operatively linked to the
CC oligonucleotide. The method is useful for preventing or treating a
CC respiratory or lung disease, which involves administering to the airways
CC of a subject an effective amount of an inhibitor. The oligonucleotide is
CC useful for production of a medicament for the prevention and/or treatment
CC of a respiratory or lung disease. The respiratory or lung disease is
CC chosen from airway inflammation, allergy(ies), asthma, impeded
CC respiration, cystic fibrosis (CF), chronic obstructive pulmonary diseases
CC (COPD), allergic rhinitis (AR), acute respiratory distress syndrome
CC (ARDS), pulmonary hypertension, lung inflammation, bronchitis, airway
CC obstruction. The present sequence represents a concatenamer of
CC oligonucleotides of the invention.
XX
SQ Sequence 4157 BP; 750 A; 1181 C; 1250 G; 976 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 12; Length 4157;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGACATGTCC 20
DB 2031 GGACATGTCCGACATGTCC 2050
RESULT 42
AD560926/C
ID AD560926 standard; DNA; 4157 BP.
XX
AC AD560926;
XX
DT 06-MAY-2004 (first entry)
XX
DE Concatemer of PDB4A oligonucleotides.
XX
KW Interleukin; IL-4 receptor; IL-5 receptor; lung disease;
KW airway inflammation; allergy; asthma; impeded respiration;
KW cystic fibrosis; acute respiratory distress syndrome;
KW pulmonary hypertension; lung inflammation; bronchitis; ds.
XX
OS Homo sapiens.
XX
PN WO2004011613-A2.
XX
PD 05-FEB-2004.
XX
PF 25-JUL-2003; 2003WO-US023509.
XX
PR 29-JUL-2002; 2002US-0399076P.
XX
PA (EPIC-) EPIGENESIS PHARM INC.
XX
PI NYCE JM, Tang L, Sandrasagra A, Aguilar D, Miller S;
PI Shahabuddin S, Lu H, Cong H;
XX
DR WPI; 2004-203534/19.
XX
PT Novel single or multiple target oligonucleotide anti-sense to e.g.
PT initiation codons and introns of respiratory disease-relevant genes e.g.,
PT CCR1, RANTES, MCP4, useful for prophylaxis or treating respiratory
PT disease e.g., asthma.
XX
PS Claim 2; SEQ ID NO 1782; 85pp; English.
XX
CC The present invention relates to an oligonucleotide anti-sense to e.g.,
CC initiation codon, coding region with 2-10 nucleotides of 5'-end and 3'-

CC end of nucleic acid target comprising gene(s) chosen from e.g.
CC interleukin (IL)-4 receptor, IL-5 receptor or salts of the
CC oligonucleotide and optionally surfactant operatively linked to the
CC oligonucleotide. The method is useful for preventing or treating a
CC respiratory or lung disease, which involves administering to the airways
CC of a subject an effective amount of an inhibitor. The oligonucleotide is
CC useful for production of a medicament for the prevention and/or treatment
CC of a respiratory or lung disease. The respiratory or lung disease is
CC chosen from airway inflammation, allergy(ies), asthma, impeded
CC respiration, cystic fibrosis (CF), chronic obstructive pulmonary diseases
CC (COPD), allergic rhinitis (AR), acute respiratory distress syndrome
CC (ARDS), pulmonary hypertension, lung inflammation, bronchitis, airway
CC obstruction. The present sequence represents a concatenamer of
CC oligonucleotides of the invention.
XX
SQ Sequence 4157 BP; 750 A; 1181 C; 1250 G; 976 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 12; Length 4157;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGACATGTCC 20
DB 2050 GGACATGTCCGACATGTCC 2031
RESULT 43
AD046415
ID AD046415 standard; DNA; 4167 BP.
XX
AC AD046415;
XX
DT 15-JUL-2004 (first entry)
XX
DE Human oligonucleotide #1781.
XX
KW Human; ss; interleukin-4 receptor; IL-4; interleukin-5 receptor; IL-5;
KW CCR1; CCR3; Botaxin-1; RANTES; MCP4; CD23; ICAM; VCAM; triptase a;
KW triptase b; PDB4 A; PDB4 B; PDB4 C; PDB4 D; respiratory disease;
KW lung disease; hyper-responsiveness; adenostine; adenostine A receptor;
KW asthma; lung allergy; inflammation; inflammatory disease;
KW airway inflammation; allergy; impeded respiration; cystic fibrosis; CF;
KW chronic obstructive pulmonary disease; COPD; allergic rhinitis;
KW acute respiratory distress syndrome; pulmonary hypertension;
KW lung inflammation; bronchitis; airway obstruction; bronchoconstriction.
XX
OS Homo sapiens.
XX
PN US2004049022-A1.
XX
PD 11-MAR-2004.
XX
PF 25-JUL-2003; 2003US-00627930.
XX
PR 23-APR-2002; 2002WO-US013135.
XX
PR 23-APR-2002; 2002WO-US013143.
XX
PA (NYCE/) NYCE J W.
PA (SAND/) SANDRASAGRA A.
PA (TANG/) TANG L.
PA (AGUI/) AGUILAR D.
PA (MILL/) MILLER S.
PA (SHAH/) SHAHABUDDIN S.
PA (LUH/) LU H.
PA (CONG/) CONG H.
XX
PI NYCE JM, Sandrasagra A, Tang L, Aguilar D, Miller S;
PI Shahabuddin S, Lu H, Cong H;
XX
DR WPI; 2004-293804/27.
XX
PT Novel single or multiple target oligonucleotide anti-sense to e.g.
PT initiation codon, intron of respiratory disease-relevant gene e.g. CCR1,

PT RANTES, MCP4, useful for prophylaxis or treating respiratory disease e.g.
 PT asthma.

Claim 2; SEQ ID NO 1782; 174pp; English.

CC The invention relates to oligonucleotides anti-sense to an initiation
 CC codon, coding region, 5' or 3' intron-exon junction, intron or region
 CC with 2-10 nucleotides of the 5'-end or 3'-end of a nucleic acid target
 CC chosen from a gene encoding interleukin (IL)-4 receptor, interleukin (IL)
 CC -5 receptor, CCR1, CCR3, Botaxin-1, RANTES, MCP4, CD23, ICAM, VCAM,
 CC tryptase a, tryptase b, PDE4 A, PDE4 B, PDE4 C or PDE4 D. The invention
 CC also relates to a method of screening a candidate compound that binds to
 CC one or more nucleic acid target(s) or expressed product(s), for the
 CC prevention and/or treatment of a respiratory or lung disease. The
 CC oligonucleotides are useful for reducing or inhibiting expression of a
 CC gene or mRNA encoding interleukin-4 receptor, interleukin-5 receptor,
 CC CCR1, CCR3, Botaxin-1, RANTES, MCP4, CD23, ICAM, VCAM, tryptase a,
 CC tryptase b, PDE4 A, PDE4 B, PDE4 C, or PDE4 D. The oligonucleotides are
 CC useful for preventing or treating a respiratory or lung disease. The
 CC respiratory or lung disease is associated with hyper-responsiveness to
 CC and/or increased levels of, adenosine and/or levels of adenosine A
 CC receptor(s), and/or asthma and/or lung allergies associated with
 CC inflammation or an inflammatory disease. The respiratory or lung disease
 CC is chosen from airway inflammation, allergy, asthma, impeded respiration,
 CC cystic fibrosis (CF), chronic obstructive pulmonary disease (COPD),
 CC allergic rhinitis, acute respiratory distress syndrome, pulmonary
 CC hyperextension, lung inflammation, bronchitis, airway obstruction or
 CC bronchoconstriction. This sequence represents an oligonucleotide of the
 CC invention.

SO Sequence 4167 BP; 750 A; 1185 C; 1250 G; 982 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 12; Length 4167;
 Best Local Similarity 90.0%; Pred. NO. 2.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGACATGCCCGGCGCATGTCC 20
 DB 2041 GGACATGTCCGACATGTCC 2060

RESULT 44
 ADO46415/C
 ID ADO46415 standard; DNA; 4167 BP.

XX ADO46415;

XX 15-JUL-2004 (first entry)

DE Human oligonucleotide #1781.

XX Human; ss; interleukin-4 receptor; IL-4; interleukin-5 receptor; IL-5;
 KW CCR1; CCR3; Botaxin-1; RANTES; MCP4; CD23; ICAM; VCAM; tryptase a;
 KW tryptase b; PDE4 A; PDE4 B; PDE4 C; PDE4 D; respiratory disease;
 KW lung disease; hyper-responsiveness; adenosine; adenosine A receptor;
 KW asthma; lung allergy; inflammation; inflammatory disease;
 KW airway inflammation; allergy; impeded respiration; cystic fibrosis; CF;
 KW chronic obstructive pulmonary disease; COPD; allergic rhinitis;
 KW acute respiratory distress syndrome; pulmonary hypertension;
 KW lung inflammation; bronchitis; airway obstruction; bronchoconstriction.

XX Homo sapiens.

XX US2004049022-A1.

XX 11-MAR-2004.

XX 25-JUL-2003; 2003US-00627930.

XX 23-APR-2002; 2002WO-US013135.

XX 23-APR-2002; 2002WO-US013143.
 XX (NYCE/) NYCE J W.

PA (SAND/) SANDRASAGRA A.
 PA (TANG/) TANG L.
 PA (AGUI/) AGUIAR D.
 PA (MILL/) MILLER S.
 PA (SHAH/) SHAHABUDDIN S.
 PA (LUHH/) LU H.
 PA (CONG/) CONG H.

PI Nye JW, Sandrasagra A, Tang L, Aguiar D, Miller S;
 PI Shahabuddin S, Lu H, Cong H;

DR WPI; 2004-293804/27.

PT Novel single or multiple target oligonucleotide anti-sense to e.g.
 PT initiation codon, intron of respiratory disease-relevant gene e.g. CCR1,
 PT RANTES, MCP4, useful for prophylaxis or treating respiratory disease e.g.
 PT asthma.

PS Claim 2; SEQ ID NO 1782; 174pp; English.

CC The invention relates to oligonucleotides anti-sense to an initiation
 CC codon, coding region, 5' or 3' intron-exon junction, intron or region
 CC with 2-10 nucleotides of the 5'-end or 3'-end of a nucleic acid target
 CC chosen from a gene encoding interleukin (IL)-4 receptor, interleukin (IL)
 CC -5 receptor, CCR1, CCR3, Botaxin-1, RANTES, MCP4, CD23, ICAM, VCAM,
 CC tryptase a, tryptase b, PDE4 A, PDE4 B, PDE4 C or PDE4 D. The invention
 CC also relates to a method of screening a candidate compound that binds to
 CC one or more nucleic acid target(s) or expressed product(s), for the
 CC prevention and/or treatment of a respiratory or lung disease. The
 CC oligonucleotides are useful for reducing or inhibiting expression of a
 CC gene or mRNA encoding interleukin-4 receptor, interleukin-5 receptor,
 CC CCR1, CCR3, Botaxin-1, RANTES, MCP4, CD23, ICAM, VCAM, tryptase a,
 CC tryptase b, PDE4 A, PDE4 B, PDE4 C, or PDE4 D. The oligonucleotides are
 CC useful for preventing or treating a respiratory or lung disease. The
 CC respiratory or lung disease is associated with hyper-responsiveness to
 CC and/or increased levels of, adenosine and/or levels of adenosine A
 CC receptor(s), and/or asthma and/or lung allergies associated with
 CC inflammation or an inflammatory disease. The respiratory or lung disease
 CC is chosen from airway inflammation, allergy, asthma, impeded respiration,
 CC cystic fibrosis (CF), chronic obstructive pulmonary disease (COPD),
 CC allergic rhinitis, acute respiratory distress syndrome, pulmonary
 CC hyperextension, lung inflammation, bronchitis, airway obstruction or
 CC bronchoconstriction. This sequence represents an oligonucleotide of the
 CC invention.

SO Sequence 4167 BP; 750 A; 1185 C; 1250 G; 982 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 12; Length 4167;
 Best Local Similarity 90.0%; Pred. NO. 2.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGACATGCCCGGCGCATGTCC 20
 DB 2060 GGACATGTCCGACATGTCC 2041

RESULT 45
 AAA88896
 ID AAA88896 standard; cDNA; 9391 BP.

XX AAA88896;

XX 05-MAR-2001 (first entry)

DE Mouse laminin 2 alpha-2 chain cDNA.

XX Laminin 2; mouse; nerve regeneration; angiogenic; cell adhesion;
 KW degenerative muscle disorder; muscular dystrophy; cell therapy; ss.
 XX Mus musculus.

XX Key Location/Qualifiers
 FT mat_peptide 1..9252

```

PT      /tag= a
XX      WO20006730-A2.
PN      09-NOV-2000.
XX      28-APR-2000; 2000WO-US011378.
XX      30-APR-1999; 99US-0131720P.
XX      15-JUN-1999; 99US-0139198P.
XX      12-JUL-1999; 99US-0143289P.
XX      24-SEP-1999; 99US-0155945P.
XX      (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX      Yurchenco P;
PI      WPI; 2000-687537/67.
XX      P-PSDB; AAB19796.
DR      Claim 4; Page 156-168; 305pp; English.
XX      The present sequence is that of cDNA encoding the mature alpha-2 chain
XX      (see AAB19796) of mouse laminin 2. Laminin 2 is composed of alpha-2 (400
XX      kDa), beta-1 (100 kDa) and gamma-1 (100 kDa) chains. It is thought to be
XX      specifically required for stabilizing myotubes during skeletal muscle
XX      development, and for preventing apoptosis. Genetic defects in human
XX      laminin 2 structure or expression are associated with a major type of
XX      congenital muscular dystrophy. Laminin 2 is also thought to be important
XX      in Schwann cell/basal lamina interactions. The invention provides laminin
XX      2 alpha-2, beta-1 and gamma-1 chain polypeptides (see AAB19791-806) and
XX      the polynucleotides encoding them (see AAB88891-906), methods for making
XX      recombinant laminin 2, cells that express recombinant laminin 2, and
XX      methods for using purified laminin 2 for research and therapeutic
XX      purposes including peripheral nerve regeneration, treatment of
XX      degenerative muscle disorders, angiogenesis regulation, promoting cell
XX      attachment and migration, ex vivo cell therapy, improving the take of
XX      grafts, improving the biocompatibility of medical devices and preparing
XX      improved culture devices and media
XX      SO      Sequence 9391 BP; 2653 A; 2153 C; 2334 G; 2241 T; 0 U; 0 Other;
Query Match      84.0%; Score 16.8; DB 3; Length 9391;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 GGACATGCCCGGCGATGCC 20
DB      793 GGTGATGCCCGGCTTGTCC 812

RESULT 46
ID      AAB88896 standard; cDNA; 9391 BP.
XX      AAB88896;
AC      05-MAR-2001 (first entry)
XX      Mouse laminin 2 alpha-2 chain cDNA.
XX      Laminin 2; mouse; nerve regeneration; angiogenic; cell adhesion;
XX      degenerative muscle disorder; muscular dystrophy; cell therapy; ss.
XX      Mus musculus.
XX      Key      Location/Qualifiers
XX      mat_peptide 1..9252
XX      FT      /tag= a

```

```

XX      WO20006730-A2.
PN      09-NOV-2000.
XX      28-APR-2000; 2000WO-US011378.
XX      30-APR-1999; 99US-0131720P.
XX      15-JUN-1999; 99US-0139198P.
XX      12-JUL-1999; 99US-0143289P.
XX      24-SEP-1999; 99US-0155945P.
XX      (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX      Yurchenco P;
PI      WPI; 2000-687537/67.
XX      P-PSDB; AAB19796.
DR      Claim 4; Page 156-168; 305pp; English.
XX      The present sequence is that of cDNA encoding the mature alpha-2 chain
XX      (see AAB19796) of mouse laminin 2. Laminin 2 is composed of alpha-2 (400
XX      kDa), beta-1 (100 kDa) and gamma-1 (100 kDa) chains. It is thought to be
XX      specifically required for stabilizing myotubes during skeletal muscle
XX      development, and for preventing apoptosis. Genetic defects in human
XX      laminin 2 structure or expression are associated with a major type of
XX      congenital muscular dystrophy. Laminin 2 is also thought to be important
XX      in Schwann cell/basal lamina interactions. The invention provides laminin
XX      2 alpha-2, beta-1 and gamma-1 chain polypeptides (see AAB19791-806) and
XX      the polynucleotides encoding them (see AAB88891-906), methods for making
XX      recombinant laminin 2, cells that express recombinant laminin 2, and
XX      methods for using purified laminin 2 for research and therapeutic
XX      purposes including peripheral nerve regeneration, treatment of
XX      degenerative muscle disorders, angiogenesis regulation, promoting cell
XX      attachment and migration, ex vivo cell therapy, improving the take of
XX      grafts, improving the biocompatibility of medical devices and preparing
XX      improved culture devices and media
XX      SO      Sequence 9391 BP; 2663 A; 2153 C; 2334 G; 2241 T; 0 U; 0 Other;
Query Match      84.0%; Score 16.8; DB 3; Length 9391;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 GGACATGCCCGGCGATGCC 20
DB      812 GGACATGCCCGGCGATGCC 793

RESULT 47
ID      AAB88895 standard; cDNA; 9511 BP.
XX      AAB88895;
AC      05-MAR-2001 (first entry)
XX      Mouse laminin 2 alpha-2 chain cDNA.
XX      Laminin 2; mouse; nerve regeneration; angiogenic; cell adhesion;
XX      degenerative muscle disorder; muscular dystrophy; cell therapy; ss.
XX      Mus musculus.
XX      Key      Location/Qualifiers
XX      CDS      55..9375
XX      FT      /tag= a
XX      sig_peptide 55..120

```

```

FT mat_peptide      /*tag= b
FT                  121..9372
FT                  /*tag= c
XX
XX MO200066730-A2.
XX
XX 09-NOV-2000.
XX
XX 28-APR-2000; 2000WO-US011378.
XX
XX 30-APR-1999; 99US-0131720P.
XX 15-JUN-1999; 99US-0139198P.
XX 12-JUL-1999; 99US-0143289P.
XX 24-SEP-1999; 99US-0155945P.
XX
XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
XX Yurchenco P;
XX
XX WPI; 2000-687537/67.
XX P-PSDB; AAB19795.
XX
XX Purified laminin 2 protein, useful for research and therapeutic purposes
XX including peripheral nerve regeneration, treatment of degenerative muscle
XX disorders, angiogenesis regulation, and ex vivo cell therapy.
XX
XX Claim 4; Page 133-146; 305pp; English.
XX
XX The present sequence is that of cDNA encoding the alpha-2 chain (see
XX AAB19795) of mouse laminin 2. Laminin 2 is composed of alpha-2 (400 kDa),
XX beta-1 (100 kDa) and gamma-1 (100 kDa) chains. It is thought to be
XX specifically required for stabilizing myotubes during skeletal muscle
XX development, and for preventing apoptosis. Genetic defects in its
XX structure or expression are associated with a major type of congenital
XX muscular dystrophy in humans. Laminin 2 is also thought to be important
XX in Schwann cell/basal lamina interactions. The invention provides laminin
XX 2 alpha-2, beta-1 and gamma-1 chain polypeptides (see AAB19791-806) and
XX recombinant laminin 2, cells that express recombinant laminin 2, and
XX methods for using purified laminin 2 for research and therapeutic
XX purposes including peripheral nerve regeneration, treatment of
XX degenerative muscle disorders, angiogenesis regulation, promoting cell
XX attachment and migration, ex vivo cell therapy, improving the take of
XX grafts, improving the biocompatibility of medical devices and preparing
XX improved culture devices and media
XX
XX Sequence 9511 BP; 2679 A; 2194 C; 2375 G; 2263 T; 0 U; 0 Other;
XX
XX Query Match      84.0%; Score 16.8; DB 3; Length 9511;
XX Best Local Similarity 90.0%; Pred. No. 2.5e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY      1 GGACATGCCCGGCGCATGTCC 20
XX      |||||
XX Db      913 GGTCATGCCCGGCGCTGTCC 932

```

```

FT CDS      55..9375
FT          /*tag= a
FT sig_peptide 55..120
FT          /*tag= b
FT mat_peptide 121..9372
FT          /*tag= c
XX
XX MO200066730-A2.
XX
XX 09-NOV-2000.
XX
XX 28-APR-2000; 2000WO-US011378.
XX
XX 30-APR-1999; 99US-0131720P.
XX 15-JUN-1999; 99US-0139198P.
XX 12-JUL-1999; 99US-0143289P.
XX 24-SEP-1999; 99US-0155945P.
XX
XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
XX Yurchenco P;
XX
XX WPI; 2000-687537/67.
XX P-PSDB; AAB19795.
XX
XX Purified laminin 2 protein, useful for research and therapeutic purposes
XX including peripheral nerve regeneration, treatment of degenerative muscle
XX disorders, angiogenesis regulation, and ex vivo cell therapy.
XX
XX Claim 4; Page 133-146; 305pp; English.
XX
XX The present sequence is that of cDNA encoding the alpha-2 chain (see
XX AAB19795) of mouse laminin 2. Laminin 2 is composed of alpha-2 (400 kDa),
XX beta-1 (100 kDa) and gamma-1 (100 kDa) chains. It is thought to be
XX specifically required for stabilizing myotubes during skeletal muscle
XX development, and for preventing apoptosis. Genetic defects in its
XX structure or expression are associated with a major type of congenital
XX muscular dystrophy in humans. Laminin 2 is also thought to be important
XX in Schwann cell/basal lamina interactions. The invention provides laminin
XX 2 alpha-2, beta-1 and gamma-1 chain polypeptides (see AAB19791-806) and
XX recombinant laminin 2, cells that express recombinant laminin 2, and
XX methods for using purified laminin 2 for research and therapeutic
XX purposes including peripheral nerve regeneration, treatment of
XX degenerative muscle disorders, angiogenesis regulation, promoting cell
XX attachment and migration, ex vivo cell therapy, improving the take of
XX grafts, improving the biocompatibility of medical devices and preparing
XX improved culture devices and media
XX
XX Sequence 9511 BP; 2679 A; 2194 C; 2375 G; 2263 T; 0 U; 0 Other;
XX
XX Query Match      84.0%; Score 16.8; DB 3; Length 9511;
XX Best Local Similarity 90.0%; Pred. No. 2.5e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY      1 GGACATGCCCGGCGCATGTCC 20
XX      |||||
XX Db      932 GGACAAAGCCCGGCGCATGACC 913

```

```

RESULT 48
AAA88895/c
ID AAA88895 standard; cDNA; 9511 BP.
XX
XX AAA88895;
XX
XX 05-MAR-2001 (first entry)
XX
XX Mouse laminin 2 alpha-2 chain cDNA.
XX
XX Laminin 2; mouse; nerve regeneration; angiogenic; cell adhesion;
XX degenerative muscle disorder; muscular dystrophy; cell therapy; ss.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
FH

```

```

RESULT 49
ADA02555
ID ADA02555 standard; DNA; 21565 BP.
XX
XX ADA02555;
XX
XX 06-NOV-2003 (first entry)
XX
XX Mouse lymphoma associated gene, SEQ ID NO:1073.
XX
XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
XX prostate; lymphoma; leukemia; cytostatic; gene therapy; drug screening;
XX gene; ds.
XX

```

```

XX Mus sp.
OS
XX
XX WO2003057146-A2.
PN
XX
XX 17-JUL-2003.
PD
XX
XX 26-DEC-2002; 2002WO-US041414.
PF
XX
XX 26-DEC-2001; 2001US-00035832.
PR
XX
XX (SAGR-) SAGRES DISCOVERY.
PA
XX
XX Morris DW;
PI
XX
XX WPI; 2003-587068/55.
DR
XX
XX New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
PS
XX Claim 1; SEQ ID NO 1073; 245bp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 21565 BP; 5397 A; 5271 C; 5466 G; 5371 T; 0 U; 60 Other;
Query Match 84.0%; Score 16.8; DB 9; Length 21565;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGCGCATGTCC 20
DB 9378 GGACATGCCCGGCGCATGTCC 9378

```

```

RESULT 50
ADA02555/C
ID ADA02555 standard; DNA; 21565 BP.
XX
XX ADA02555;
AC
XX
XX 06-NOV-2003 (first entry)
DT
XX
XX Mouse lyfe carcinoma associated gene, SEQ ID NO:1073.
DE
XX
XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KM prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX
XX Mus sp.
OS
XX
XX WO2003057146-A2.
PN

```

```

XX
XX 17-JUL-2003.
PD
XX
XX 26-DEC-2002; 2002WO-US041414.
PF
XX
XX 26-DEC-2001; 2001US-00035832.
PR
XX
XX (SAGR-) SAGRES DISCOVERY.
PA
XX
XX Morris DW;
PI
XX
XX WPI; 2003-587068/55.
DR
XX
XX New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
PS
XX Claim 1; SEQ ID NO 1073; 245bp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 21565 BP; 5397 A; 5271 C; 5466 G; 5371 T; 0 U; 60 Other;
Query Match 84.0%; Score 16.8; DB 9; Length 21565;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGACATGCCCGGCGCATGTCC 20
DB 9378 GGACATGCCCGGCGCATGTCC 9378

```

Search completed: January 14, 2005, 16:14:16
Job time : 360.737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 14:47:55 ; Search time 71.0526 Seconds
(without alignments)
200.074 Million cell updates/sec

Title: US-09-578-453-2
Perfect score: 20
Sequence: 1 GCACATCCCGCGCATGTC 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database: Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	2	US-08-904-901-162
2	20	100.0	20	2	US-08-904-901-162
3	20	100.0	20	3	US-08-249-730-162
4	20	100.0	20	3	US-08-249-730-162
5	20	100.0	20	4	US-09-249-247-162
6	20	100.0	20	4	US-09-249-247-162
7	20	100.0	26	3	US-08-767-942A-40
8	20	100.0	26	3	US-08-767-942A-40
9	20	100.0	27	3	US-08-767-942A-41
10	20	100.0	27	3	US-08-767-942A-41
11	16.8	84.0	30	1	US-08-347-792-18
12	16.8	84.0	30	1	US-08-347-792-18
13	16.8	84.0	30	1	US-08-431-357-18
14	16.8	84.0	30	1	US-08-431-357-18
15	16.8	84.0	30	1	US-08-697-221-29
16	16.8	84.0	30	2	US-08-697-221-29
17	16.8	84.0	30	2	US-08-392-542-31
18	16.8	84.0	30	3	US-08-392-542-31
19	16.8	84.0	30	3	US-08-894-327-31
20	16.8	84.0	30	3	US-08-894-327-31
21	16.8	84.0	30	4	US-09-685-027-31
22	16.8	84.0	30	4	US-09-685-027-31
23	16.8	84.0	30	4	US-09-829-922-31
24	16.8	84.0	30	4	US-09-829-922-31
25	16.8	84.0	30	5	PCT-US95-15353-18
26	16.8	84.0	30	5	PCT-US95-15353-18
27	16.8	84.0	9391	4	US-09-562-702A-11

c	28	16.8	84.0	9391	4	US-09-562-702A-11	Sequence 11, Appl
c	29	16.8	84.0	9511	4	US-09-562-702A-9	Sequence 9, Appl
c	30	16.8	84.0	9511	4	US-09-562-702A-9	Sequence 9, Appl
c	31	15.4	77.0	473	4	US-09-312-283C-355	Sequence 355, App
c	32	15.4	77.0	473	4	US-09-312-283C-355	Sequence 355, App
c	33	15.2	76.0	20	2	US-08-657-828A-3	Sequence 3, Appl
c	34	15.2	76.0	20	2	US-08-657-828A-3	Sequence 3, Appl
c	35	15.2	76.0	20	3	US-09-260-420-3	Sequence 3, Appl
c	36	15.2	76.0	20	3	US-09-260-420-3	Sequence 3, Appl
c	37	15.2	76.0	291	4	US-09-513-999C-22333	Sequence 22333, A
c	38	15.2	76.0	291	4	US-09-513-999C-22333	Sequence 22333, A
c	39	15.2	76.0	1251	4	US-09-828-000-1	Sequence 1, Appl
c	40	15.2	76.0	1251	4	US-09-828-000-1	Sequence 1, Appl
c	41	15.2	76.0	1920	4	US-09-906-393A-35	Sequence 35, Appl
c	42	15.2	76.0	1920	4	US-09-906-393A-35	Sequence 35, Appl
c	43	15.2	76.0	1958	4	US-09-702-327-3	Sequence 3, Appl
c	44	15.2	76.0	1958	4	US-09-702-327-3	Sequence 3, Appl
c	45	15.2	76.0	1958	4	US-09-460-309-3	Sequence 3, Appl
c	46	15.2	76.0	6942	2	US-08-460-309-3	Sequence 3, Appl
c	47	15.2	76.0	6942	2	US-08-125-077-3	Sequence 3, Appl
c	48	15.2	76.0	6942	2	US-08-125-077-3	Sequence 3, Appl
c	49	15.2	76.0	9419	4	US-09-562-702A-7	Sequence 7, Appl
c	50	15.2	76.0	9419	4	US-09-562-702A-7	Sequence 7, Appl
c	51	15.2	76.0	9420	4	US-09-562-702A-3	Sequence 3, Appl
c	52	15.2	76.0	9420	4	US-09-562-702A-3	Sequence 3, Appl
c	53	15.2	76.0	9534	4	US-09-562-702A-5	Sequence 5, Appl
c	54	15.2	76.0	9534	4	US-09-562-702A-5	Sequence 5, Appl
c	55	15.2	76.0	9534	4	US-09-561-709B-8	Sequence 8, Appl
c	56	15.2	76.0	9534	4	US-09-561-709B-8	Sequence 8, Appl
c	57	15.2	76.0	9534	4	US-09-917-254-35	Sequence 35, Appl
c	58	15.2	76.0	9535	4	US-09-917-254-35	Sequence 35, Appl
c	59	15.2	76.0	9535	4	US-09-562-702A-1	Sequence 1, Appl
c	60	15.2	76.0	9535	4	US-09-562-702A-1	Sequence 1, Appl
c	61	15.2	76.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
c	62	15.2	76.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
c	63	15.2	76.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
c	64	15.2	76.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
c	65	15	75.0	1743	4	US-09-252-991A-13408	Sequence 13408, A
c	66	15	75.0	1743	4	US-09-252-991A-13408	Sequence 13408, A
c	67	15	75.0	2199	4	US-09-252-991A-13503	Sequence 13503, A
c	68	15	75.0	2199	4	US-09-252-991A-13503	Sequence 13503, A
c	69	15	75.0	2433	4	US-09-252-991A-13950	Sequence 13950, A
c	70	15	75.0	2433	4	US-09-252-991A-13950	Sequence 13950, A
c	71	14.8	74.0	49	2	US-08-299-074A-31	Sequence 31, Appl
c	72	14.8	74.0	49	2	US-08-299-074A-31	Sequence 31, Appl
c	73	14.8	74.0	49	3	US-09-399-773-31	Sequence 31, Appl
c	74	14.8	74.0	49	3	US-09-399-773-31	Sequence 31, Appl
c	75	14.8	74.0	516	4	US-09-252-991A-4742	Sequence 4742, Ap
c	76	14.8	74.0	516	4	US-09-252-991A-4742	Sequence 4742, Ap
c	77	14.8	74.0	762	4	US-09-252-991A-4766	Sequence 4766, Ap
c	78	14.8	74.0	762	4	US-09-252-991A-4766	Sequence 4766, Ap
c	79	14.8	74.0	811	4	US-09-040-229B-5	Sequence 5, Appl
c	80	14.8	74.0	811	4	US-09-040-229B-5	Sequence 5, Appl
c	81	14.8	74.0	1027	4	US-09-684-405-8	Sequence 8, Appl
c	82	14.8	74.0	1027	4	US-09-684-405-8	Sequence 8, Appl
c	83	14.8	74.0	1781	4	US-09-649-747A-20	Sequence 20, Appl
c	84	14.8	74.0	1781	4	US-09-649-747A-20	Sequence 20, Appl
c	85	14.8	74.0	1949	4	US-09-649-747A-12	Sequence 12, Appl
c	86	14.8	74.0	1949	4	US-09-649-747A-12	Sequence 12, Appl
c	87	14.8	74.0	87350	3	US-08-781-891-79	Sequence 79, Appl
c	88	14.8	74.0	87350	3	US-08-781-891-79	Sequence 79, Appl
c	89	14.8	74.0	87350	4	US-09-618-166-79	Sequence 79, Appl
c	90	14.8	74.0	87350	4	US-09-618-166-79	Sequence 79, Appl
c	91	14.8	74.0	87543	4	US-09-791-211-3	Sequence 3, Appl
c	92	14.8	74.0	87543	4	US-09-791-211-3	Sequence 3, Appl
c	93	14.4	72.0	5066	4	US-09-824-574-1	Sequence 1, Appl
c	94	14.4	72.0	5066	4	US-09-824-574-1	Sequence 1, Appl
c	95	14.2	71.0	303	4	US-09-252-991A-7045	Sequence 7045, Ap
c	96	14.2	71.0	303	4	US-09-252-991A-7045	Sequence 7045, Ap
c	97	14.2	71.0	393	4	US-09-252-991A-4966	Sequence 4966, Ap
c	98	14.2	71.0	393	4	US-09-252-991A-4966	Sequence 4966, Ap
c	99	14.2	71.0	568	4	US-09-513-999C-1079	Sequence 1079, Ap
c	100	14.2	71.0	568	4	US-09-513-999C-1079	Sequence 1079, Ap

101	14.2	71.0	693	4	US-09-252-991A-13194	Sequence 13194, A	c 174	13.8	69.0	578	4	US-09-604-287A-150	Sequence 150, App
c 102	14.2	71.0	693	4	US-09-252-991A-13194	Sequence 13194, A	c 175	13.8	69.0	578	4	US-09-285-480-150	Sequence 150, App
c 103	14.2	71.0	726	4	US-09-252-991A-4997	Sequence 4997, Ap	c 176	13.8	69.0	578	4	US-09-285-480-150	Sequence 150, App
c 104	14.2	71.0	726	4	US-09-252-991A-4997	Sequence 4997, Ap	c 177	13.8	69.0	578	4	US-09-834-759-150	Sequence 150, App
c 105	14.2	71.0	759	4	US-09-252-991A-13714	Sequence 13714, A	c 178	13.8	69.0	578	4	US-09-834-759-150	Sequence 150, App
c 106	14.2	71.0	759	4	US-09-252-991A-13714	Sequence 13714, A	c 179	13.8	69.0	578	4	US-09-590-751A-150	Sequence 150, App
c 107	14.2	71.0	798	4	US-09-252-991A-13429	Sequence 13429, A	c 180	13.8	69.0	578	4	US-09-590-751A-150	Sequence 150, App
c 108	14.2	71.0	798	4	US-09-252-991A-13429	Sequence 13429, A	c 181	13.8	69.0	649	3	US-09-328-111-314	Sequence 111, App
c 109	14.2	71.0	816	4	US-09-252-991A-4951	Sequence 4951, Ap	c 182	13.8	69.0	799	3	US-09-328-111-314	Sequence 111, App
c 110	14.2	71.0	816	4	US-09-252-991A-4951	Sequence 4951, Ap	c 183	13.8	69.0	799	3	US-08-998-416-420	Sequence 416, App
c 111	14.2	71.0	816	4	US-09-252-991A-13914	Sequence 13914, A	c 184	13.8	69.0	799	3	US-08-998-416-420	Sequence 416, App
c 112	14.2	71.0	816	4	US-09-252-991A-13914	Sequence 13914, A	c 185	13.8	69.0	1001	4	US-09-671-317-347	Sequence 317, App
c 113	14.2	71.0	966	4	US-09-252-991A-12761	Sequence 12761, A	c 186	13.8	69.0	1001	4	US-09-671-317-347	Sequence 317, App
c 114	14.2	71.0	966	4	US-09-252-991A-12761	Sequence 12761, A	c 187	13.8	69.0	1556	3	US-09-026-958-1	Sequence 1, App1
c 115	14.2	71.0	1125	4	US-09-252-991A-11202	Sequence 11202, A	c 188	13.8	69.0	1556	3	US-09-026-958-1	Sequence 1, App1
c 116	14.2	71.0	1125	4	US-09-252-991A-11202	Sequence 11202, A	c 189	13.8	69.0	1810	4	US-09-800-729-73	Sequence 73, App1
c 117	14.2	71.0	1134	4	US-09-252-991A-12761	Sequence 12761, A	c 190	13.8	69.0	1810	4	US-09-800-729-73	Sequence 73, App1
c 118	14.2	71.0	1134	4	US-09-252-991A-12761	Sequence 12761, A	c 191	13.8	69.0	1811	4	US-09-800-729-77	Sequence 77, App1
c 119	14.2	71.0	1140	1	US-08-314-172A-13	Sequence 13, App1	c 192	13.8	69.0	1811	4	US-09-800-729-77	Sequence 77, App1
c 120	14.2	71.0	1140	1	US-08-314-172A-13	Sequence 13, App1	c 193	13.8	69.0	1964	3	US-09-434-613-2	Sequence 2, App1
c 121	14.2	71.0	1254	4	US-09-252-991A-10642	Sequence 10642, A	c 194	13.8	69.0	1964	3	US-09-434-613-2	Sequence 2, App1
c 122	14.2	71.0	1254	4	US-09-252-991A-10642	Sequence 10642, A	c 195	13.8	69.0	1964	4	US-09-963-908-2	Sequence 2, App1
c 123	14.2	71.0	1308	4	US-09-252-991A-12614	Sequence 12614, A	c 196	13.8	69.0	1964	4	US-09-963-908-2	Sequence 2, App1
c 124	14.2	71.0	1308	4	US-09-252-991A-12614	Sequence 12614, A	c 197	13.8	69.0	1985	4	US-09-620-312D-812	Sequence 812, App
c 125	14.2	71.0	1488	4	US-09-252-991A-4938	Sequence 4938, Ap	c 198	13.8	69.0	1985	4	US-09-620-312D-812	Sequence 812, App
c 126	14.2	71.0	1488	4	US-09-252-991A-4938	Sequence 4938, Ap	c 199	13.8	69.0	4768	4	US-09-526-193A-16	Sequence 16, App1
c 127	14.2	71.0	1668	4	US-09-252-991A-13705	Sequence 13705, A	c 200	13.8	69.0	4768	4	US-09-526-193A-16	Sequence 16, App1
c 128	14.2	71.0	1668	4	US-09-252-991A-13705	Sequence 13705, A	c 201	13.6	68.0	20	3	US-09-210-748A-6	Sequence 6, App1
c 129	14.2	71.0	1803	3	US-09-369-364A-20	Sequence 20, App1	c 202	13.6	68.0	20	3	US-09-210-748A-6	Sequence 6, App1
c 130	14.2	71.0	1803	3	US-09-369-364A-20	Sequence 20, App1	c 203	13.6	68.0	20	4	US-09-939-581A-6	Sequence 6, App1
c 131	14.2	71.0	1947	4	US-09-252-991A-7111	Sequence 7111, Ap	c 204	13.6	68.0	132	4	US-09-513-999C-18033	Sequence 18033, A
c 132	14.2	71.0	1947	4	US-09-252-991A-7111	Sequence 7111, Ap	c 205	13.6	68.0	132	4	US-09-513-999C-18033	Sequence 18033, A
c 133	14.2	71.0	2001	4	US-09-213-888-26	Sequence 26, App1	c 206	13.6	68.0	132	4	US-09-513-999C-18033	Sequence 18033, A
c 134	14.2	71.0	2001	4	US-09-213-888-26	Sequence 26, App1	c 207	13.6	68.0	293	4	US-09-313-294A-5620	Sequence 5620, App
c 135	14.2	71.0	2001	4	US-09-328-877D-26	Sequence 26, App1	c 208	13.6	68.0	293	4	US-09-313-294A-5620	Sequence 5620, App
c 136	14.2	71.0	2001	4	US-09-328-877D-26	Sequence 26, App1	c 209	13.6	68.0	341	4	US-09-621-976-18936	Sequence 18936, A
c 137	14.2	71.0	2010	4	US-09-213-888-24	Sequence 24, App1	c 210	13.6	68.0	341	4	US-09-621-976-18936	Sequence 18936, A
c 138	14.2	71.0	2010	4	US-09-213-888-24	Sequence 24, App1	c 211	13.6	68.0	366	4	US-09-513-999C-13574	Sequence 13574, A
c 139	14.2	71.0	2010	4	US-09-328-877D-24	Sequence 24, App1	c 212	13.6	68.0	366	4	US-09-513-999C-13574	Sequence 13574, A
c 140	14.2	71.0	2010	4	US-09-328-877D-24	Sequence 24, App1	c 213	13.6	68.0	368	4	US-09-404-879A-269	Sequence 269, App
c 141	14.2	71.0	2169	4	US-09-252-991A-10432	Sequence 10432, A	c 214	13.6	68.0	368	4	US-09-404-879A-269	Sequence 269, App
c 142	14.2	71.0	2169	4	US-09-252-991A-10432	Sequence 10432, A	c 215	13.6	68.0	368	4	US-09-338-933-269	Sequence 269, App
c 143	14.2	71.0	2979	4	US-09-252-991A-10524	Sequence 10524, A	c 216	13.6	68.0	368	4	US-09-338-933-269	Sequence 269, App
c 144	14.2	71.0	2979	4	US-09-252-991A-10524	Sequence 10524, A	c 217	13.6	68.0	368	4	US-09-215-681-269	Sequence 269, App
c 145	14.2	71.0	3550	4	US-09-213-888-1	Sequence 1, App1	c 218	13.6	68.0	368	4	US-09-215-681-269	Sequence 269, App
c 146	14.2	71.0	3550	4	US-09-213-888-1	Sequence 1, App1	c 219	13.6	68.0	368	4	US-09-215-681-269	Sequence 269, App
c 147	14.2	71.0	3550	4	US-09-328-877D-1	Sequence 1, App1	c 220	13.6	68.0	368	4	US-09-215-681-269	Sequence 269, App
c 148	14.2	71.0	3550	4	US-09-328-877D-1	Sequence 1, App1	c 221	13.6	68.0	368	4	US-09-216-003A-269	Sequence 269, App
c 149	14.2	71.0	3559	4	US-09-800-729-53	Sequence 53, App1	c 222	13.6	68.0	368	4	US-09-667-857-269	Sequence 269, App
c 150	14.2	71.0	3559	4	US-09-800-729-53	Sequence 53, App1	c 223	13.6	68.0	429	4	US-09-667-857-269	Sequence 269, App
c 151	14.2	71.0	5720	4	US-09-800-729-18	Sequence 18, App1	c 224	13.6	68.0	429	4	US-09-107-532A-160	Sequence 160, App
c 152	14.2	71.0	5720	4	US-09-800-729-18	Sequence 18, App1	c 225	13.6	68.0	447	4	US-09-107-532A-160	Sequence 160, App
c 153	14.2	71.0	193303	4	US-09-497-855A-37	Sequence 37, App1	c 226	13.6	68.0	447	4	US-09-621-976-9361	Sequence 9361, Ap
c 154	14.2	71.0	193303	4	US-09-497-855A-37	Sequence 37, App1	c 227	13.6	68.0	541	4	US-09-621-976-18562	Sequence 18562, A
c 155	14.2	71.0	193303	4	US-09-497-855A-44	Sequence 44, App1	c 228	13.6	68.0	541	4	US-09-621-976-18562	Sequence 18562, A
c 156	14.2	71.0	193303	4	US-09-497-855A-44	Sequence 44, App1	c 229	13.6	68.0	798	4	US-09-270-767-12368	Sequence 12368, A
c 157	14.2	71.0	536165	4	US-09-214-808-1	Sequence 1, App1	c 230	13.6	68.0	798	4	US-09-270-767-12368	Sequence 12368, A
c 158	14.2	71.0	536165	4	US-09-214-808-1	Sequence 1, App1	c 231	13.6	68.0	801	4	US-09-252-991A-4545	Sequence 4545, Ap
c 159	14	70.0	35881	4	US-08-311-731A-127	Sequence 127, App	c 232	13.6	68.0	801	4	US-09-252-991A-4545	Sequence 4545, Ap
c 160	14	70.0	35881	4	US-08-311-731A-127	Sequence 127, App	c 233	13.6	68.0	834	4	US-09-252-991A-348	Sequence 348, App
c 161	13.8	69.0	4335	4	US-09-621-976-16911	Sequence 16911, A	c 234	13.6	68.0	834	4	US-09-252-991A-348	Sequence 348, App
c 162	13.8	69.0	4335	4	US-09-621-976-16911	Sequence 16911, A	c 235	13.6	68.0	843	4	US-09-292-858B-5	Sequence 5, App1
c 163	13.8	69.0	578	3	US-09-222-575-150	Sequence 150, App	c 236	13.6	68.0	843	4	US-09-292-858B-5	Sequence 5, App1
c 164	13.8	69.0	578	3	US-09-222-575-150	Sequence 150, App	c 237	13.6	68.0	901	3	US-09-328-111-2	Sequence 2, App1
c 165	13.8	69.0	578	4	US-09-389-681-150	Sequence 150, App	c 238	13.6	68.0	901	3	US-09-328-111-2	Sequence 2, App1
c 166	13.8	69.0	578	4	US-09-389-681-150	Sequence 150, App	c 239	13.6	68.0	945	4	US-09-107-532A-3562	Sequence 3562, Ap
c 167	13.8	69.0	578	4	US-09-620-405B-150	Sequence 150, App	c 240	13.6	68.0	945	4	US-09-107-532A-3562	Sequence 3562, Ap
c 168	13.8	69.0	578	4	US-09-620-405B-150	Sequence 150, App	c 241	13.6	68.0	969	4	US-09-252-991A-4308	Sequence 4308, Ap
c 169	13.8	69.0	578	4	US-09-339-338-150	Sequence 150, App	c 242	13.6	68.0	969	4	US-09-252-991A-4308	Sequence 4308, Ap
c 170	13.8	69.0	578	4	US-09-339-338-150	Sequence 150, App	c 243	13.6	68.0	999	4	US-09-724-797-17	Sequence 17, App1
c 171	13.8	69.0	578	4	US-09-433-826B-150	Sequence 150, App	c 244	13.6	68.0	999	4	US-09-724-797-17	Sequence 17, App1
c 172	13.8	69.0	578	4	US-09-433-826B-150	Sequence 150, App	c 245	13.6	68.0	1047	4	US-09-252-991A-13520	Sequence 13520, A
c 173	13.8	69.0	578	4	US-09-604-287A-150	Sequence 150, App	c 246	13.6	68.0	1047	4	US-09-252-991A-13520	Sequence 13520, A

247	13.6	68.0	1080	4	US-09-891-641-63	Sequence 63, Appl	C 320	13.6	68.0	8878	1	US-08-759-444-2	Sequence 2, Appl
C 248	13.6	68.0	1080	4	US-09-891-641-63	Sequence 63, Appl	C 321	13.6	68.0	9880	3	US-08-680-897-1	Sequence 1, Appl
C 249	13.6	68.0	1200	4	US-09-620-312D-1095	Sequence 1095, Ap	C 322	13.6	68.0	9880	3	US-08-680-897-1	Sequence 1, Appl
C 250	13.6	68.0	1300	4	US-09-620-312D-1095	Sequence 1095, Ap	C 323	13.6	68.0	1187	4	US-09-422-936-61	Sequence 61, Appl
C 251	13.6	68.0	1392	4	US-09-252-991A-12249	Sequence 12249, A	C 324	13.6	68.0	1187	4	US-09-422-936-61	Sequence 61, Appl
C 252	13.6	68.0	1392	4	US-09-252-991A-12249	Sequence 12249, A	C 325	13.6	68.0	1187	4	US-09-422-936-61	Sequence 61, Appl
C 253	13.6	68.0	1401	4	US-09-252-991A-120	Sequence 320, App	C 326	13.6	68.0	11863	4	US-09-814-915A-83	Sequence 83, Appl
C 254	13.6	68.0	1401	4	US-09-252-991A-120	Sequence 320, App	C 327	13.6	68.0	11863	4	US-09-814-915A-83	Sequence 83, Appl
C 255	13.6	68.0	1407	4	US-09-252-991A-12361	Sequence 12361, A	C 328	13.6	68.0	15602	4	US-09-844-634-17	Sequence 17, Appl
C 256	13.6	68.0	1407	4	US-09-252-991A-12361	Sequence 12361, A	C 329	13.6	68.0	15602	4	US-09-844-634-17	Sequence 17, Appl
C 257	13.6	68.0	1530	4	US-09-252-991A-12109	Sequence 12109, A	C 330	13.6	68.0	37950	3	US-09-338-907-183	Sequence 183, App
C 258	13.6	68.0	1530	4	US-09-252-991A-12109	Sequence 12109, A	C 331	13.6	68.0	37950	3	US-09-338-907-183	Sequence 183, App
C 259	13.6	68.0	1596	4	US-09-252-991A-12109	Sequence 12109, A	C 332	13.6	68.0	37950	3	US-09-218-207-183	Sequence 183, App
C 260	13.6	68.0	1596	4	US-09-252-991A-12109	Sequence 12109, A	C 333	13.6	68.0	37950	3	US-09-218-207-183	Sequence 183, App
C 261	13.6	68.0	1620	3	US-09-360-197-3	Sequence 9478, Ap	C 334	13.6	68.0	55298	4	US-09-491-356C-1	Sequence 1, Appl
C 262	13.6	68.0	1620	3	US-09-360-197-3	Sequence 9478, Ap	C 335	13.6	68.0	55298	4	US-09-491-356C-1	Sequence 1, Appl
C 263	13.6	68.0	1650	4	US-09-252-991A-9352	Sequence 9352, Ap	C 336	13.4	67.0	30	3	US-08-648-506-1	Sequence 1, Appl
C 264	13.6	68.0	1650	4	US-09-252-991A-9352	Sequence 9352, Ap	C 337	13.4	67.0	30	3	US-08-648-506-1	Sequence 1, Appl
C 265	13.6	68.0	1778	4	US-09-799-451-790	Sequence 790, App	C 338	13.4	67.0	30	3	US-08-648-506-1	Sequence 1, Appl
C 266	13.6	68.0	1778	4	US-09-799-451-790	Sequence 790, App	C 339	13.4	67.0	366	4	US-09-277-355-1	Sequence 1, Appl
C 267	13.6	68.0	2089	1	US-08-552-142A-1	Sequence 1, Appl	C 340	13.4	67.0	366	4	US-09-277-355-1	Sequence 1, Appl
C 268	13.6	68.0	2089	1	US-08-552-142A-1	Sequence 1, Appl	C 341	13.4	67.0	417	4	US-09-569-852B-4	Sequence 4, Appl
C 269	13.6	68.0	2089	1	US-08-910-973-1	Sequence 1, Appl	C 342	13.4	67.0	417	4	US-09-569-852B-4	Sequence 4, Appl
C 270	13.6	68.0	2089	1	US-08-910-973-1	Sequence 1, Appl	C 343	13.4	67.0	417	4	US-09-569-852B-4	Sequence 4, Appl
C 271	13.6	68.0	2089	1	US-08-910-973-1	Sequence 1, Appl	C 344	13.4	67.0	543	4	US-09-328-352-957	Sequence 3957, Ap
C 272	13.6	68.0	2089	4	US-09-499-227-1	Sequence 1, Appl	C 345	13.4	67.0	543	4	US-09-328-352-957	Sequence 3957, Ap
C 273	13.6	68.0	2089	5	PCT-US95-05741-1	Sequence 1, Appl	C 346	13.4	67.0	603	3	US-08-998-416-1152	Sequence 1152, Ap
C 274	13.6	68.0	2089	5	PCT-US95-05741-1	Sequence 1, Appl	C 347	13.4	67.0	1026	4	US-08-998-416-1152	Sequence 1152, Ap
C 275	13.6	68.0	2091	4	US-09-252-991A-12312	Sequence 12312, A	C 348	13.4	67.0	1026	4	US-08-998-416-1152	Sequence 1152, Ap
C 276	13.6	68.0	2091	4	US-09-252-991A-12312	Sequence 12312, A	C 349	13.4	67.0	1027	4	US-09-252-991A-3491	Sequence 3491, Ap
C 277	13.6	68.0	2224	3	US-08-477-347-2	Sequence 2, Appl	C 350	13.4	67.0	1142	3	US-09-674-741-9	Sequence 9, Appl
C 278	13.6	68.0	2224	3	US-08-477-347-2	Sequence 2, Appl	C 351	13.4	67.0	1142	3	US-09-674-741-9	Sequence 9, Appl
C 279	13.6	68.0	2224	3	US-08-477-347-2	Sequence 2, Appl	C 352	13.4	67.0	1142	3	US-09-674-741-9	Sequence 9, Appl
C 280	13.6	68.0	2224	3	US-08-477-347-2	Sequence 2, Appl	C 353	13.4	67.0	1142	3	US-09-674-741-9	Sequence 9, Appl
C 281	13.6	68.0	2224	3	US-08-477-347-2	Sequence 2, Appl	C 354	13.4	67.0	1142	3	US-09-674-741-9	Sequence 9, Appl
C 282	13.6	68.0	2224	4	US-09-800-909-1	Sequence 1, Appl	C 355	13.4	67.0	1172	4	US-09-311-021-149	Sequence 149, App
C 283	13.6	68.0	2224	4	US-09-800-909-1	Sequence 1, Appl	C 356	13.4	67.0	1172	4	US-09-311-021-149	Sequence 149, App
C 284	13.6	68.0	2224	4	US-09-800-909-1	Sequence 1, Appl	C 357	13.4	67.0	1302	4	US-09-252-991A-3473	Sequence 3473, Ap
C 285	13.6	68.0	2266	2	US-09-800-908-2	Sequence 2, Appl	C 358	13.4	67.0	1332	4	US-09-252-991A-3473	Sequence 3473, Ap
C 286	13.6	68.0	2266	2	US-09-800-908-2	Sequence 2, Appl	C 359	13.4	67.0	1332	4	US-09-252-991A-3473	Sequence 3473, Ap
C 287	13.6	68.0	2266	2	US-08-724-394A-18	Sequence 18, Appl	C 360	13.4	67.0	1722	4	US-09-252-991A-3485	Sequence 3485, Ap
C 288	13.6	68.0	2431	4	US-09-221-017B-351	Sequence 351, App	C 361	13.4	67.0	1722	4	US-09-252-991A-3485	Sequence 3485, Ap
C 289	13.6	68.0	2490	4	US-09-221-017B-351	Sequence 351, App	C 362	13.4	67.0	1762	4	US-09-799-451-122	Sequence 622, App
C 290	13.6	68.0	2490	4	US-09-221-017B-351	Sequence 351, App	C 363	13.4	67.0	1904	4	US-09-799-451-122	Sequence 622, App
C 291	13.6	68.0	2613	4	US-09-856-937A-1	Sequence 10579, A	C 364	13.4	67.0	1904	4	US-09-799-451-122	Sequence 622, App
C 292	13.6	68.0	2613	4	US-09-856-937A-1	Sequence 10579, A	C 365	13.4	67.0	1904	4	US-09-799-451-122	Sequence 622, App
C 293	13.6	68.0	2613	4	US-09-856-937A-1	Sequence 10579, A	C 366	13.4	67.0	2907	4	US-10-140-002-99	Sequence 99, Appl
C 294	13.6	68.0	2818	4	US-09-491-356C-4	Sequence 4, Appl	C 367	13.4	67.0	2907	4	US-10-140-002-99	Sequence 99, Appl
C 295	13.6	68.0	2818	4	US-09-491-356C-4	Sequence 4, Appl	C 368	13.4	67.0	4517	3	US-09-023-655-1053	Sequence 1053, Ap
C 296	13.6	68.0	2974	1	US-08-290-978A-4	Sequence 4, Appl	C 369	13.4	67.0	4517	3	US-09-023-655-1053	Sequence 1053, Ap
C 297	13.6	68.0	2974	1	US-08-290-978A-4	Sequence 4, Appl	C 370	13.4	67.0	4517	3	US-09-023-655-1053	Sequence 1053, Ap
C 298	13.6	68.0	2974	2	US-08-780-869-4	Sequence 4, Appl	C 371	13.4	67.0	4517	4	US-09-140-804-9	Sequence 9, Appl
C 299	13.6	68.0	3323	4	US-10-140-002-167	Sequence 167, App	C 372	13.4	67.0	4517	4	US-09-140-804-9	Sequence 9, Appl
C 300	13.6	68.0	3323	4	US-10-140-002-167	Sequence 167, App	C 373	13.4	67.0	4517	4	US-09-140-804-9	Sequence 9, Appl
C 301	13.6	68.0	3384	4	US-09-252-991A-10278	Sequence 10278, A	C 374	13.4	67.0	4517	4	US-09-140-804-9	Sequence 9, Appl
C 302	13.6	68.0	3384	4	US-09-252-991A-10278	Sequence 10278, A	C 375	13.4	67.0	4517	4	US-09-140-804-9	Sequence 9, Appl
C 303	13.6	68.0	3403	4	US-09-023-655-737	Sequence 737, App	C 376	13.4	67.0	4517	4	US-09-140-804-9	Sequence 9, Appl
C 304	13.6	68.0	3403	4	US-09-023-655-737	Sequence 737, App	C 377	13.4	67.0	4517	4	US-09-140-804-9	Sequence 9, Appl
C 305	13.6	68.0	3492	4	US-09-023-655-1219	Sequence 1219, App	C 378	13.4	67.0	4517	4	US-09-140-804-9	Sequence 9, Appl
C 306	13.6	68.0	3492	4	US-09-023-655-1219	Sequence 1219, App	C 379	13.4	67.0	5526	3	US-08-907-146-21	Sequence 21, Appl
C 307	13.6	68.0	3683	4	US-09-844-634-3	Sequence 3, Appl	C 380	13.4	67.0	5526	3	US-08-907-146-21	Sequence 21, Appl
C 308	13.6	68.0	3683	4	US-09-844-634-3	Sequence 3, Appl	C 381	13.4	67.0	5526	3	US-08-907-146-21	Sequence 21, Appl
C 309	13.6	68.0	3683	4	US-09-844-634-3	Sequence 3, Appl	C 382	13.2	66.0	20	4	US-08-751-359-21	Sequence 21, Appl
C 310	13.6	68.0	3683	4	US-09-844-634-3	Sequence 3, Appl	C 383	13.2	66.0	20	4	US-08-751-359-21	Sequence 21, Appl
C 311	13.6	68.0	4233	4	US-09-491-356C-5	Sequence 5, Appl	C 384	13.2	66.0	137	3	US-08-998-416-465	Sequence 465, App
C 312	13.6	68.0	4233	4	US-09-491-356C-5	Sequence 5, Appl	C 385	13.2	66.0	137	3	US-08-998-416-465	Sequence 465, App
C 313	13.6	68.0	4436	4	US-09-491-356C-6	Sequence 6, Appl	C 386	13.2	66.0	138	3	US-08-998-416-465	Sequence 465, App
C 314	13.6	68.0	4436	4	US-09-491-356C-6	Sequence 6, Appl	C 387	13.2	66.0	224	4	US-09-513-999C-14238	Sequence 14238, A
C 315	13.6	68.0	5140	3	US-09-333-214-2	Sequence 2, Appl	C 388	13.2	66.0	224	4	US-09-513-999C-14238	Sequence 14238, A
C 316	13.6	68.0	5140	3	US-09-333-214-2	Sequence 2, Appl	C 389	13.2	66.0	263	4	US-09-636-215-682	Sequence 682, App
C 317	13.6	68.0	7244	4	US-09-774-528-143	Sequence 143, App	C 390	13.2	66.0	263	4	US-09-636-215-682	Sequence 682, App
C 318	13.6	68.0	7244	4	US-09-774-528-143	Sequence 143, App	C 391	13.2	66.0	263	4	US-09-636-215-682	Sequence 682, App
C 319	13.6	68.0	8878	1	US-08-759-444-2	Sequence 2, Appl	C 392	13.2	66.0	263	4	US-09-636-215-682	Sequence 682, App

393	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
394	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
395	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
396	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
397	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
398	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
399	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
400	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
401	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
402	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
403	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
404	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
405	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
406	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
407	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
408	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
409	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
410	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
411	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
412	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
413	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
414	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
415	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
416	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
417	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
418	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
419	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
420	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
421	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
422	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
423	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
424	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
425	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
426	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
427	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
428	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
429	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
430	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
431	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
432	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
433	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
434	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
435	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
436	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
437	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
438	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
439	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
440	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
441	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
442	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
443	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
444	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
445	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
446	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
447	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
448	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
449	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
450	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
451	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
452	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
453	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
454	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
455	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
456	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
457	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
458	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
459	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
460	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
461	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
462	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
463	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
464	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App
465	13.2	66.0	263	4	US-09-679-426-682	Sequence 682, App

466	13.2	66.0	1649	4	US-09-484-970B-72	Sequence 72, App
467	13.2	66.0	1654	4	US-09-221-017B-326	Sequence 326, App
468	13.2	66.0	1654	4	US-09-221-017B-326	Sequence 326, App
469	13.2	66.0	1666	3	US-09-360-197-5	Sequence 5, App
470	13.2	66.0	1666	3	US-09-360-197-5	Sequence 5, App
471	13.2	66.0	1674	3	US-09-360-197-5	Sequence 5, App
472	13.2	66.0	1674	3	US-09-360-197-5	Sequence 5, App
473	13.2	66.0	1785	4	US-09-620-312D-593	Sequence 593, App
474	13.2	66.0	1785	4	US-09-489-039A-5174	Sequence 5174, App
475	13.2	66.0	1908	4	US-09-489-039A-5174	Sequence 5174, App
476	13.2	66.0	1908	4	US-09-489-039A-5174	Sequence 5174, App
477	13.2	66.0	2176	4	US-09-252-991A-1914	Sequence 1914, App
478	13.2	66.0	2176	4	US-09-843-250-8	Sequence 1914, App
479	13.2	66.0	2400	1	US-09-843-250-8	Sequence 8, App
480	13.2	66.0	2400	1	US-09-843-250-8	Sequence 8, App
481	13.2	66.0	2400	1	US-09-843-250-8	Sequence 8, App
482	13.2	66.0	2400	2	US-09-843-250-8	Sequence 8, App
483	13.2	66.0	2400	2	US-09-843-250-8	Sequence 8, App
484	13.2	66.0	2400	3	US-09-843-250-8	Sequence 8, App
485	13.2	66.0	2400	3	US-09-843-250-8	Sequence 8, App
486	13.2	66.0	2400	4	US-09-843-250-8	Sequence 8, App
487	13.2	66.0	2571	4	US-09-843-250-8	Sequence 8, App
488	13.2	66.0	2571	4	US-09-843-250-8	Sequence 8, App
489	13.2	66.0	4355	4	US-09-843-250-8	Sequence 8, App
490	13.2	66.0	4355	4	US-09-843-250-8	Sequence 8, App
491	13.2	66.0	4848	3	US-09-843-250-8	Sequence 8, App
492	13.2	66.0	4848	3	US-09-843-250-8	Sequence 8, App
493	13.2	66.0	4848	3	US-09-843-250-8	Sequence 8, App
494	13.2	66.0	4848	3	US-09-843-250-8	Sequence 8, App
495	13.2	66.0	4848	3	US-09-843-250-8	Sequence 8, App
496	13.2	66.0	4848	3	US-09-843-250-8	Sequence 8, App
497	13.2	66.0	5510	4	US-08-956-171E-165	Sequence 6, App
498	13.2	66.0	5510	4	US-08-956-171E-165	Sequence 6, App
499	13.2	66.0	5510	4	US-08-956-171E-165	Sequence 6, App
500	13.2	66.0	5510	4	US-08-956-171E-165	Sequence 6, App

RESULT 1
 US-08-904-901-162
 ; Sequence 162, Application US/08904901
 ; Patent No. 5998383
 GENERAL INFORMATION:
 APPLICANT: Wright, Jim A.
 APPLICANT: Young, Alping H.
 TITLE OF INVENTION: ANTITUMOR ANTISENSE SEQUENCES DIRECTED
 AGAINST RIBONUCLEOTIDE REDUCTASE
 NUMBER OF SEQUENCES: 163
 CORRESPONDENCE ADDRESS:
 ADDRESSER: KOHN & ASSOCIATES
 STREET: 30500 No. 5998383Western Hwy. Suite 410
 CITY: Farmington Hills
 STATE: Michigan
 COUNTRY: US
 ZIP: 48334
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/904,901
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Kohn, Kenneth I.
 REGISTRATION NUMBER: 30,955
 REFERENCE/DOCKET NUMBER: 0227.00004
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (248) 539-5050
 TELEFAX: (248) 539-5055

INFORMATION FOR SEQ ID NO: 162:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Primer"
US-08-904-901-162

Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20
DB 1 GGACATGCCGGGCGCATGTCC 20

RESULT 2

US-08-904-901-162/c
Sequence 162, Application US/08904901
Patent No. 5998383
GENERAL INFORMATION:
APPLICANT: Wright, Jim A.
TITLE OF INVENTION: ANTITUMOR ANTISENSE SEQUENCES DIRECTED
TITLE OF INVENTION: AGAINST RIBONUCLEOTIDE REDUCTASE
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: KOHN & ASSOCIATES
STREET: 30500 No. 5998383thwestern Hwy. Suite 410
City: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,901
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0227,00004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 162:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Primer"
US-08-904-901-162

Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20
DB 20 GGACATGCCGGGCGCATGTCC 1

RESULT 3
US-09-249-730-162

Sequence 162, Application US/09249730
Patent No. 6121000
GENERAL INFORMATION:
APPLICANT: WRIGHT, Jim A.
TITLE OF INVENTION: Antitumor Antisense Sequences Directed Against R1 and
TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase
FILE REFERENCE: 032396-040
CURRENT APPLICATION NUMBER: US/09/249,730
CURRENT FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 220
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 162
LENGTH: 20
TYPE: DNA
ORGANISM: Human
US-09-249-730-162

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20
DB 1 GGACATGCCGGGCGCATGTCC 20

RESULT 4

US-09-249-730-162/c
Sequence 162, Application US/09249730
Patent No. 6121000
GENERAL INFORMATION:
APPLICANT: WRIGHT, Jim A.
TITLE OF INVENTION: Antitumor Antisense Sequences Directed Against R1 and
TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase
FILE REFERENCE: 032396-040
CURRENT APPLICATION NUMBER: US/09/249,730
CURRENT FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 220
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 162
LENGTH: 20
TYPE: DNA
ORGANISM: Human
US-09-249-730-162

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20
DB 20 GGACATGCCGGGCGCATGTCC 1

RESULT 5

US-09-249-730-162
Sequence 162, Application US/09249247
Patent No. 6593305
GENERAL INFORMATION:
APPLICANT: WRIGHT, Jim A.
TITLE OF INVENTION: Antitumor Antisense Sequences Directed Against R1 and
TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase
FILE REFERENCE: 032396-023
CURRENT APPLICATION NUMBER: US/09/249,247
CURRENT FILING DATE: 1999-02-11
EARLIER APPLICATION NUMBER: US 60/023,040
EARLIER FILING DATE: 1996-08-02
EARLIER APPLICATION NUMBER: US 60/039,959
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: US 08/904,901

EARLIER FILING DATE: 1997-08-01
NUMBER OF SEQ ID NOS: 220
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 162
LENGTH: 20
TYPE: DNA
ORGANISM: Human
US-09-249-247-162

Query Match 100.0%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
DB 1 GGACATGCCCGGCATGTCC 20

RESULT 6
US-09-249-247-162/c
Sequence 162, Application US/09249247
Patent No. 6593305
GENERAL INFORMATION:

APPLICANT: WRIGHT, Jim A.
APPLICANT: YOUNG, Aiping H.
TITLE OF INVENTION: Anticancer Antisense Sequences Directed Against R1 and
TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase
FILE REFERENCE: 032396-023
CURRENT FILING DATE: 1999-02-11
EARLIER APPLICATION NUMBER: US 60/023,040
EARLIER FILING DATE: 1996-08-02
EARLIER APPLICATION NUMBER: US 60/039,959
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: US 08/904,901
EARLIER FILING DATE: 1997-08-01
NUMBER OF SEQ ID NOS: 220
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 162
LENGTH: 20
TYPE: DNA
ORGANISM: Human
US-09-249-247-162

Query Match 100.0%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
DB 20 GGACATGCCCGGCATGTCC 1

RESULT 7
US-08-767-942A-40
Sequence 40, Application US/08767942A
Patent No. 6068982
GENERAL INFORMATION:

APPLICANT: Rolfe, Mark
APPLICANT: Chiu, M. Isabel
APPLICANT: Berlin, Vivian
APPLICANT: Damagnez, Veronique
APPLICANT: Draetta, Giulio
APPLICANT: Guillaume, Coctarel
TITLE OF INVENTION: UBIQUITIN CONUGATING ENZYMES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,942A
FILING DATE: 17-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-029.04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-767-942A-40

Query Match 100.0%; Score 20; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
DB 6 GGACATGCCCGGCATGTCC 25

RESULT 8
US-08-767-942A-40/c
Sequence 40, Application US/08767942A
Patent No. 6068982
GENERAL INFORMATION:

APPLICANT: Rolfe, Mark
APPLICANT: Chiu, M. Isabel
APPLICANT: Berlin, Vivian
APPLICANT: Damagnez, Veronique
APPLICANT: Draetta, Giulio
APPLICANT: Guillaume, Coctarel
TITLE OF INVENTION: UBIQUITIN CONUGATING ENZYMES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,942A
FILING DATE: 17-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-029.04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-767-942A-40

Query Match 100.0%; Score 20; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 25 GGACATGCCCGGCGCATGTCC 6

RESULT 9

US-08-767-942A-41
Sequence 41, Application US/08767942A

PATENT No. 6068982
GENERAL INFORMATION:
APPLICANT: Rolfe, Mark
APPLICANT: Chiu, M. Isabel
APPLICANT: Berlin, Vivian
APPLICANT: Damagnez, Veronique
APPLICANT: Draetta, Giulio
APPLICANT: Guillaume, Cottarel
TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSER: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/767,942A
APPLICATION NUMBER: US/08/767,942A
FILING DATE: 17-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-029.04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULAR TYPE: DNA (genomic)
US-08-767-942A-41

Query Match 100.0%; Score 20; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 7 GGACATGCCCGGCGCATGTCC 26

RESULT 10

US-08-767-942A-41/C
Sequence 41, Application US/08767942A

PATENT No. 6068982
GENERAL INFORMATION:
APPLICANT: Rolfe, Mark
APPLICANT: Chiu, M. Isabel

APPLICANT: Berlin, Vivian
APPLICANT: Damagnez, Veronique
APPLICANT: Draetta, Giulio
APPLICANT: Guillaume, Cottarel
TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSER: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,942A
FILING DATE: 17-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-029.04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-767-942A-41

Query Match 100.0%; Score 20; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 26 GGACATGCCCGGCGCATGTCC 7

RESULT 11
US-08-347-792-18
Sequence 18, Application US/08347792

PATENT No. 5573925
GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos D.
TITLE OF INVENTION: p53 Proteins With Altered
TITLE OF INVENTION: Tetramerization Domains
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSER: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/347,792
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST58USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-347-792-18

Query Match 84.0%; Score 16.8; DB 1; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20
|||
DB 3 GGGCATGTCCGGGCATGTCC 22

RESULT 12
US-08-347-792-18/c
Sequence 18, Application US/08347792
Patent No. 5573925
GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos D.
TITLE OF INVENTION: p53 Proteins With Altered
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSER: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/347,792
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST58USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-347-792-18

Query Match 84.0%; Score 16.8; DB 1; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20
|||
DB 22 GGACATGCCCGGCATGTCC 3

RESULT 13
US-08-431-357-18
Sequence 18, Application US/08431357
Patent No. 5721340

GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos D.
TITLE OF INVENTION: p53 Proteins With Altered
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSER: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/347,792
FILING DATE: 28-NOV-1994

ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST58USA

TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-431-357-18

Query Match 84.0%; Score 16.8; DB 1; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20
|||
DB 3 GGGCATGTCCGGGCATGTCC 22

RESULT 14
US-08-431-357-18/c
Sequence 18, Application US/08431357
Patent No. 5721340
GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos D.
TITLE OF INVENTION: p53 Proteins With Altered
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSER: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/347,792
FILING DATE: 28-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST58USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-431-357-18

Query Match 84.0%; Score 16.8; DB 1; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 20
DB 22 GGACATGCCCGGCGCATGCC 3

RESULT 15
US-08-697-221-29
Sequence 29, Application US/08697221
Patent No. 5847083
GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos D.
TITLE OF INVENTION: Modified p53 Constructs and Uses
NUMBER OF INVENTION: Therefore
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/697,221
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,802
FILING DATE: 22-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kodroff, Cathy A.
REGISTRATION NUMBER: 33,980
REFERENCE/DOCKET NUMBER: WST64AUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
US-08-697-221-29

Query Match 84.0%; Score 16.8; DB 2; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 20
DB 3 GGCGATGTCGGGCGCATGTC 22

RESULT 16
US-08-697-221-29/c
Sequence 29, Application US/08697221
Patent No. 5847083
GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos D.
TITLE OF INVENTION: Modified p53 Constructs and Uses
NUMBER OF INVENTION: Therefore
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/697,221
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,802
FILING DATE: 22-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kodroff, Cathy A.
REGISTRATION NUMBER: 33,980
REFERENCE/DOCKET NUMBER: WST64AUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
US-08-697-221-29

Query Match 84.0%; Score 16.8; DB 2; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 20
DB 22 GGACATGCCCGGCGCATGCC 3

RESULT 17
US-08-392-542-31
Sequence 31, Application US/08392542
Patent No. 6169073
GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos
ATTORNEY/AGENT INFORMATION:
NAME: Hartwig, Wolfgang

TITLE OF INVENTION: Peptides nad Peptidomimetics with
TITLE OF INVENTION: Structural Similarity to Human p53 That Activate p53
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,542
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Posorske, Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 0486,48439
TELEPHONE: 202 508-9100
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-392-542-31

Query Match 84.0%; Score 16.8; DB 3; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 3 GGCGATGTCCGGCGCATGTCC 22

RESULT 18
US-08-392-542-31/c
Sequence 31, Application US/08392542
Patent No. 6169073
GENERAL INFORMATION:
APPLICANT: Halazonecis, Thanos
APPLICANT: Hartwig, Wolfgang
TITLE OF INVENTION: Peptides nad Peptidomimetics with
TITLE OF INVENTION: Structural Similarity to Human p53 That Activate p53
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,542
FILING DATE:

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Posorske, Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 0486,48439
TELEPHONE: 202 508-9100
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-392-542-31

Query Match 84.0%; Score 16.8; DB 3; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 22 GGACATGCCCGGCGCATGTCC 3

RESULT 19
US-08-894-327-31
Sequence 31, Application US/08894327
Patent No. 6245886
GENERAL INFORMATION:
APPLICANT: Halazonecis, Thanos
APPLICANT: Hartwig, Wolfgang
TITLE OF INVENTION: Peptides and peptidomimetics with
TITLE OF INVENTION: structural similarity to human p53 that activate p53
FILE REFERENCE: 2973,19998
CURRENT APPLICATION NUMBER: US/08/894,327
CURRENT FILING DATE: 1997-12-04
EARLIER APPLICATION NUMBER: pctus96/01535
EARLIER FILING DATE: 1996-02-16
EARLIER APPLICATION NUMBER: 08/392,542
EARLIER FILING DATE: 1995-02-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 30
TYPE: DNA
ORGANISM: Homo sapiens
US-08-894-327-31

Query Match 84.0%; Score 16.8; DB 3; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 3 GGCGATGTCCGGCGCATGTCC 22

RESULT 20
US-08-894-327-31/c
Sequence 31, Application US/08894327
Patent No. 6245886
GENERAL INFORMATION:
APPLICANT: Halazonecis, Thanos
APPLICANT: Hartwig, Wolfgang
TITLE OF INVENTION: Peptides and peptidomimetics with
TITLE OF INVENTION: structural similarity to human p53 that activate p53
FILE REFERENCE: 2973,19998

CURRENT APPLICATION NUMBER: US/08/894,327
CURRENT FILING DATE: 1997-12-04
EARLIER APPLICATION NUMBER: Pctus96/01535
EARLIER FILING DATE: 1996-02-16
EARLIER APPLICATION NUMBER: 08/392,542
EARLIER FILING DATE: 1995-02-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 30
TYPE: DNA
ORGANISM: Homo sapiens
US-08-894-327-31

Query Match 84.0%; Score 16.8; DB 3; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCGCGGATGTC 20
22 GGACATGCGCGGATGTC 3

RESULT 21
US-09-685-027-31
Sequence 31, Application US/09685027
Patent No. 6420118
GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos
Hartwig, Wolfgang
TITLE OF INVENTION: Peptides and Peptidomimetics with
Structural Similarity to Human p53 That Activate p53
Function
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/685,027
FILING DATE: 10-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/392,542
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Posorske, Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 0486,48439
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9100
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-685-027-31
Query Match 84.0%; Score 16.8; DB 4; Length 30;

Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCGCGGATGTC 20
Db 3 GGACATGTCGGGATGTC 22

RESULT 22
US-09-685-027-31/C
Sequence 31, Application US/09685027
Patent No. 6420118
GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos
Hartwig, Wolfgang
TITLE OF INVENTION: Peptides and Peptidomimetics with
Structural Similarity to Human p53 That Activate p53
Function
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/685,027
FILING DATE: 10-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/392,542
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Posorske, Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 0486,48439
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9100
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-685-027-31
Query Match 84.0%; Score 16.8; DB 4; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCGCGGATGTC 20
Db 22 GGACATGCGCGGATGTC 3

RESULT 23
US-09-829-922-31
Sequence 31, Application US/09829922
Patent No. 6784157
GENERAL INFORMATION:
APPLICANT: Halazonetis, Thanos
Hartwig, Wolfgang

```

? TITLE OF INVENTION: Peptides and peptidomimetics with
? TITLE OF INVENTION: structural similarity to human p53 that activate p53
? TITLE OF INVENTION:
? TITLE OF INVENTION: function
? FILE REFERENCE: 2973.19998
? CURRENT APPLICATION NUMBER: US/09/829,922
? CURRENT FILING DATE: 2001-04-11
? PRIOR APPLICATION NUMBER: 08/894,327
? PRIOR FILING DATE: 1997-12-04
? PRIOR APPLICATION NUMBER: pctus96/01535
? PRIOR FILING DATE: 1996-02-16
? PRIOR APPLICATION NUMBER: 08/392,542
? PRIOR FILING DATE: 1995-02-16
? NUMBER OF SEQ ID NOS: 35
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 31
? LENGTH: 30
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-829-922-31

```

Query Match	84.0%	Score 16.8	DB 4	Length 30
Best Local Similarity	90.0%	Pred. No. 52		
Matches 18, Conservative	0	Mismatches 2	Indels 0	Gaps 0

QY 1 GGACATGCCCCGGGCATGTC 20
2
3 GGCATGTCGGGCATGTC 22
Db

```

RESULT 24
US-09-829-922-31/c
; Sequence 31, Application US/09829922
; Patent No. 6784157
; GENERAL INFORMATION:
; APPLICANT: Halazonecis, Thamos
; APPLICANT: Hartwig, Wolfgang
; TITLE OF INVENTION: Peptides and peptidomimetics with
; TITLE OF INVENTION: structural similarity to human p53 that activate p53
; TITLE OF INVENTION: function
; FILE REFERENCE: 2973.19998
; CURRENT APPLICATION NUMBER: US/09/829,922
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 08/894,327
; PRIOR FILING DATE: 1997-12-04
; PRIOR APPLICATION NUMBER: pcus96/01535
; PRIOR FILING DATE: 1996-02-16
; PRIOR APPLICATION NUMBER: 08/392,542
; PRIOR FILING DATE: 1995-02-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-829-922-31

```

Query Match	84.0%;	Score 16.8;	DB 4;	Length 30;
Best Local Similarity	90.0%;	Pred. No. 52;		
Matches 18; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

```

QY      1 GGACATGCCCGGCATGTC 20
          |||||
Db      22 GGACATGCCCGGCATGCCC 3

```

RESULT 25
 PCT-US95-15353-18
 ; Sequence 18, Application PC/TUS9515353
 ; GENERAL INFORMATION:
 ; APPLICANT: The Wistar Institute of Anatomy
 ; APPLICANT: and Biology
 ; APPLICANT: Halazonetis, Thamos D.
 TITLE OF INVENTION: p53 Proteins With Altered

1 TITLE OF INVENTION: Polymerization Domains
2
3 NUMBER OF SEQUENCES: 46
4 CORRESPONDENCE ADDRESS:
5 ADDRESSEE: Howson and Howson
6 STREET: Spring House Corporate Cntr., PO Box 45
7 CITY: Spring House
8 STATE: Pennsylvania
9 COUNTRY: USA
10
11 ZIP: 19477
12

```

? COMPUTER READABLE FORM:
?
? MEDIUM TYPE: Floppy disk
?
? COMPUTER: IBM PC compatible
?
? OPERATING SYSTEM: PC-DOS/MS-DOS
?
? SOFTWARE: Patent In Release #1.0, Version #1.3
?
? CURRENT APPLICATION DATA:

```

CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/347,792
FILING DATE: 28-NOV-1994
PRIOR APPLICATION DATA:

;;
; FILING DATE: 28-APR-1995
; APPLICATION NUMBER: US 08/431,357
; PRIOR APPLICATION DATA:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/456,623
;; FILING DATE: 01-JUN-1995

ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31 015

REGISTRATION NUMBER: 31,215
REFERENCE NUMBER: WST58CPC7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 315-540-9206

TELEPHONE: 215-540-5200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 18
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

Query Match	84.0%;	Score 16.8;	DB 5;	Length 30;
Best Local Similarity	90.0%;	Pred. No. 52;		
Matches 18; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy 1 GGACATGCCCCGGGCATGTCC 20
|||
Db 3 GGGCATGTCCGGGCATGTCC 22

RESULT 26
PCT-US95-15353-18/c
Sequence 18, Application PC/RUS9515353
GENERAL INFORMATION:
APPLICANT: The Wistar Institute of Anatomy
and Biology
APPLICANT: Halaszetis, Thamos D.
TITLE OF INVENTION: p53 Proteins With Altered
Tetramerization Domains
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15353
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/347,792
FILING DATE: 28-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/431,357
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/456,623
FILING DATE: 01-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST58CPCPT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US95-15353-18

Query Match 84.0%; Score 16.8; DB 5; Length 30;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 22 GGACATGCCCGGCGCATGTCC 3

RESULT 27
US-09-562-702A-11
Sequence 11, Application US/09562702A
Patent No. 6632790
GENERAL INFORMATION:
APPLICANT: Yurchenco, Peter
TITLE OF INVENTION: Laminin 2 and Methods for Its Use
FILE REFERENCE: 99-274-B
CURRENT APPLICATION NUMBER: US/09/562,702A
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/155,945
PRIOR FILING DATE: 1999-09-24
PRIOR APPLICATION NUMBER: 60/143,289
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/139,198
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/131,720
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 11
LENGTH: 9391
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(9252)
US-09-562-702A-11

Query Match 84.0%; Score 16.8; DB 4; Length 9391;
Best Local Similarity 90.0%; Pred. No. 53;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20

DB 793 GGTCATGCCCGGCGCTTGTC 812

RESULT 28
US-09-562-702A-11/c
Sequence 11, Application US/09562702A
Patent No. 6632790
GENERAL INFORMATION:
APPLICANT: Yurchenco, Peter
TITLE OF INVENTION: Laminin 2 and Methods for Its Use
FILE REFERENCE: 99-274-B
CURRENT APPLICATION NUMBER: US/09/562,702A
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/155,945
PRIOR FILING DATE: 1999-09-24
PRIOR APPLICATION NUMBER: 60/143,289
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/139,198
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/131,720
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 11
LENGTH: 9391
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(9252)
US-09-562-702A-11

Query Match 84.0%; Score 16.8; DB 4; Length 9391;
Best Local Similarity 90.0%; Pred. No. 53;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 812 GGACATGCCCGGCGCATGTCC 793

RESULT 29
US-09-562-702A-9
Sequence 9, Application US/09562702A
Patent No. 6632790
GENERAL INFORMATION:
APPLICANT: Yurchenco, Peter
TITLE OF INVENTION: Laminin 2 and Methods for Its Use
FILE REFERENCE: 99-274-B
CURRENT APPLICATION NUMBER: US/09/562,702A
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/155,945
PRIOR FILING DATE: 1999-09-24
PRIOR APPLICATION NUMBER: 60/143,289
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/139,198
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/131,720
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 9
LENGTH: 9511
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (55)..(9372)
NAME/KEY: sig_peptide
LOCATION: (55)..(120)
US-09-562-702A-9

Query Match 84.0%; Score 16.8; DB 4; Length 9511;
Best Local Similarity 90.0%; Pred. No. 53;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 20
DB 913 GGTCATGCCCGGCGCTGTTC 932

RESULT 30

US-09-562-702A-9/C
Sequence 9, Application US/09562702A
Patent No. 6632790
GENERAL INFORMATION:
APPLICANT: Yurchenco, Peter
TITLE OF INVENTION: Laminin 2 and Methods for Its Use
FILE REFERENCE: 99-274-B
CURRENT APPLICATION NUMBER: US/09/562,702A
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/155,945
PRIOR FILING DATE: 1999-09-24
PRIOR APPLICATION NUMBER: 60/143,289
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/139,198
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/131,720
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 9511
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (55)..
NAME/KEY: sig peptide
LOCATION: (55)..
US-09-562-702A-9

Query Match 84.0%; Score 16.8; DB 4; Length 9511;
Best Local Similarity 90.0%; Pred. No. 53;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 20
DB 932 GGACATGCCCGGCGCATGACC 913

RESULT 31

US-09-312-283C-355
Sequence 355, Application US/09312283C
Patent No. 6573095
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Ornust, Rene
APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated from Skin Cells
FILE REFERENCE: 11000.1011c2
CURRENT APPLICATION NUMBER: US/09/312,283C
PRIOR FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 425
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 355
LENGTH: 473
TYPE: DNA
ORGANISM: Mouse
US-09-312-283C-355

Query Match 77.0%; Score 15.4; DB 4; Length 473;
Best Local Similarity 94.1%; Pred. No. 2.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATG 17
DB 367 GGACATGCCCTGGCATG 383

RESULT 32

US-09-312-283C-355/C
Sequence 355, Application US/09312283C
Patent No. 6573095
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Ornust, Rene
APPLICANT: Murison, James G.
TITLE OF INVENTION: Compositions Isolated from Skin Cells
FILE REFERENCE: 11000.1011c2
CURRENT APPLICATION NUMBER: US/09/312,283C
PRIOR FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 425
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 355
LENGTH: 473
TYPE: DNA
ORGANISM: Mouse
US-09-312-283C-355

Query Match 77.0%; Score 15.4; DB 4; Length 473;
Best Local Similarity 94.1%; Pred. No. 2.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CATGCCCGGCGCATGTC 20
DB 383 CATGCCCGGCGCATGTC 367

US-08-657-828A-3
Sequence 3, Application US/08657828A
Patent No. 5876711

GENERAL INFORMATION:
APPLICANT: Fattaey, Ali
TITLE OF INVENTION: Methods and Compositions for Determining
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Onyx Pharmaceuticals, Inc.
STREET: 3031 Research Drive
CITY: Richmond
STATE: CA
COUNTRY: USA
ZIP: 94806
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/657,828A
FILING DATE: 31-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Glotta, Gregory
REGISTRATION NUMBER: 32,028
REFERENCE/DOCKET NUMBER: ONYX1021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-262-8710

TELEFAX: 510-758-3405
INFORMATION FOR SEQ ID NO: 3
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-657-828A-3

Query Match 76.0%; Score 15.2; DB 2; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGCC 20
Db 1 RRCATGYRRRCATGYTY 20

RESULT 34
US-08-657-828A-3/c
Sequence 3, Application US/08657828A
Patent No. 5876711

GENERAL INFORMATION:
APPLICANT: Fattaei, Ali
TITLE OF INVENTION: Methods and Compositions for Determining
TITLE OF INVENTION: the Tumor Suppressor Status of Cells
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Onyx Pharmaceuticals, Inc.
STREET: 3031 Research Drive
CITY: Richmond
STATE: CA
COUNTRY: USA
ZIP: 94806
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/657,828A
FILING DATE: 31-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gioceta, Gregory
REGISTRATION NUMBER: 32,028
REFERENCE/DOCKET NUMBER: ONYX1021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-262-8710
TELEFAX: 510-758-3405
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-657-828A-3

Query Match 76.0%; Score 15.2; DB 2; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGCC 20
Db 20 RRCATGYRRRCATGYTY 1

RESULT 35
US-09-260-420-3
Sequence 3, Application US/09260420
Patent No. 6391630

GENERAL INFORMATION:
APPLICANT: Fattaei, Ali
TITLE OF INVENTION: Methods and Compositions for Determining
TITLE OF INVENTION: the Tumor Suppressor Status of Cells
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Onyx Pharmaceuticals, Inc.
STREET: 3031 Research Drive
CITY: Richmond
STATE: CA
COUNTRY: USA
ZIP: 94806
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/260,420
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/657,828
FILING DATE: 31-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gioceta, Gregory
REGISTRATION NUMBER: 32,028
REFERENCE/DOCKET NUMBER: ONYX1021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-262-8710
TELEFAX: 510-758-3405
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-260-420-3

Query Match 76.0%; Score 15.2; DB 3; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGCC 20
Db 1 RRCATGYRRRCATGYTY 20

RESULT 36
US-09-260-420-3/c
Sequence 3, Application US/09260420
Patent No. 6391630
GENERAL INFORMATION:
APPLICANT: Fattaei, Ali
TITLE OF INVENTION: Methods and Compositions for Determining
TITLE OF INVENTION: the Tumor Suppressor Status of Cells
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Onyx Pharmaceuticals, Inc.
STREET: 3031 Research Drive
CITY: Richmond
STATE: CA
COUNTRY: USA
ZIP: 94806
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/260,420
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/657,828
; FILING DATE: 31-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Giotta, Gregory
; REGISTRATION NUMBER: 32,028
; REFERENCE/DOCKET NUMBER: ONYX1021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-262-8710
; TELEFAX: 510-758-3405
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-09-260-420-3

```

```

Query Match          76.0%; Score 15.2; DB 4; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GGACATGCCCGGCGCATGTC 20
DB      20 RRCATGYRRRCATGYT 1

```

```

RESULT 37
US-09-513-999C-22333
; Sequence 22333, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 22333
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 288
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-22333

```

```

Query Match          76.0%; Score 15.2; DB 4; Length 291;
Best Local Similarity 85.0%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      1 GGACATGCCCGGCGCATGTC 20
DB      160 GGACTTCCCCGGGAGATGTC 179

```

```

RESULT 38
US-09-513-999C-22333/c
; Sequence 22333, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 22333
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 288
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-22333

```

```

Query Match          76.0%; Score 15.2; DB 4; Length 291;
Best Local Similarity 85.0%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 GGACATGCCCGGCGCATGTC 20
DB      179 GCAATCCCGGCGCAATGCC 160

```

```

RESULT 39
US-09-828-000-1
; Sequence 1, Application US/09828000
; Patent No. 6596690
; GENERAL INFORMATION:
; APPLICANT: Government of the United States of America
; TITLE OF INVENTION: Vasostratin as Marrow Protectant
; FILE REFERENCE: 4239-55414
; CURRENT APPLICATION NUMBER: US/09/828,000
; CURRENT FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Calreticulin
US-09-828-000-1

```

```

Query Match          76.0%; Score 15.2; DB 4; Length 1251;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 GGACATGCCCGGCGCATGTC 20
DB      708 GGACAAGCCCGGAGATATCC 727

```

```

RESULT 40
US-09-828-000-1/c
; Sequence 1, Application US/09828000
; Patent No. 6596690
; GENERAL INFORMATION:
; APPLICANT: Government of the United States of America
; TITLE OF INVENTION: Vasostratin as Marrow Protectant
; FILE REFERENCE: 4239-55414
; CURRENT APPLICATION NUMBER: US/09/828,000
; CURRENT FILING DATE: 2001-04-06

```


NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 1251
TYPE: DNA
ORGANISM: Calreticulin
US-09-828-000-1

Query Match 76.0%; Score 15.2; DB 4; Length 1251;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 727 GGATATGCTCGGCGCTGTCC 708

RESULT 41
US-09-906-393A-35
Sequence 35, Application US/09906393A
Patent No. 6780984
GENERAL INFORMATION:
APPLICANT: Wang, Zhou
APPLICANT: Xiao, Wuhan
TITLE OF INVENTION: METHOD OF PROGNOSING CANCER AND THE PROTEINS INVOLVED
FILE REFERENCE: 1720-1-001CIP
CURRENT APPLICATION NUMBER: US/09/906,393A
CURRENT FILING DATE: 2001-07-16
PRIOR FILING DATE: 2000-07-17
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Patentin version 3.1
SEQ ID NO 35
LENGTH: 1920
TYPE: DNA
ORGANISM: Homo sapiens
US-09-906-393A-35

Query Match 76.0%; Score 15.2; DB 4; Length 1920;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 778 GGACAGCCCGAGCATATCC 797

RESULT 42
US-09-906-393A-35/C
Sequence 35, Application US/09906393A
Patent No. 6780984
GENERAL INFORMATION:
APPLICANT: Wang, Zhou
APPLICANT: Xiao, Wuhan
TITLE OF INVENTION: METHOD OF PROGNOSING CANCER AND THE PROTEINS INVOLVED
FILE REFERENCE: 1720-1-001CIP
CURRENT APPLICATION NUMBER: US/09/906,393A
CURRENT FILING DATE: 2001-07-16
PRIOR FILING DATE: 2000-07-17
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Patentin version 3.1
SEQ ID NO 35
LENGTH: 1920
TYPE: DNA
ORGANISM: Homo sapiens
US-09-906-393A-35

Query Match 76.0%; Score 15.2; DB 4; Length 1920;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20

DB 797 GGATATGCTCGGCGCTGTCC 778

RESULT 43
US-09-702-327-3
Sequence 3, Application US/09702327
Patent No. 6426220
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Lex M. Cowsett
TITLE OF INVENTION: ANTISENSE MODULATION OF CALRETICULIN EXPRESSION
FILE REFERENCE: RTS-0097
CURRENT APPLICATION NUMBER: US/09/702,327
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 3
LENGTH: 1958
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: CDS
LOCATION: (109)..(1362)
US-09-702-327-3

Query Match 76.0%; Score 15.2; DB 4; Length 1958;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 816 GGACAGCCCGAGCATATCC 835

RESULT 44
US-09-702-327-3/C
Sequence 3, Application US/09702327
Patent No. 6426220
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Lex M. Cowsett
TITLE OF INVENTION: ANTISENSE MODULATION OF CALRETICULIN EXPRESSION
FILE REFERENCE: RTS-0097
CURRENT APPLICATION NUMBER: US/09/702,327
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 3
LENGTH: 1958
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: CDS
LOCATION: (109)..(1362)
US-09-702-327-3

Query Match 76.0%; Score 15.2; DB 4; Length 1958;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
DB 835 GGATATGCTCGGCGCTGTCC 816

RESULT 45
US-08-460-309-3
Sequence 3, Application US/08460309
Patent No. 5837496
GENERAL INFORMATION:
APPLICANT: Bngvall, Eva
APPLICANT: Leiyo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof

```

: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/460,309
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/125,077
: FILING DATE: 22-SEP-1993
: APPLICATION NUMBER: US PCT/US 94/10730
: FILING DATE: 21-SEP-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/472,319
: FILING DATE: 30-JAN-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/919,951
: FILING DATE: 27-JUL-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P-LA 9721
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001
: TELEFAX: (619) 535-8849
: INFORMATION FOR SEQ ID NO. 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6942 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-460-309-3
:
: Query Match 76.0%; Score 15.2; DB 2; Length 6942;
: Best Local Similarity 85.0%; Pred. No. 2.9e+02;
: Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0,
:
: Oy 1 GGACATGCCCGGGGATGCC 20
: Db 939 GGACAGCCTTGGCATGACC 920
:
: RESULT 47
: US-08-125-077-3
: Sequence 3, Application US/08125077
: Patent No. 5872231
: Patent No. 5872231 5840863
: GENERAL INFORMATION:
: APPLICANT: Engvall, Eva
: APPLICANT: Letivo, Ilmo
: TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
: TITLE OF INVENTION: Fragments and Uses Thereof
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Campbell and Flores
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92122
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/125,077
: FILING DATE: 22-SEP-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US PCT/US 94/10730
: FILING DATE: 21-SEP-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/472,319
: FILING DATE: 30-JAN-1990

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/919,951
; FILING DATE: 27-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9721
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6942 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-125-077-3
;
Query Match 76.0%; Score 15.2; DB 2; Length 6942;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
Db 920 GGTCATGCCAGGCTGTGTC 939

RESULT 48
US-08-125-077-3/c
; Sequence 3, Application US/08125077
; Patent No. 5872231
; Patent No. 5872231 5840863
; GENERAL INFORMATION:
; APPLICANT: Engvall, Eva
; APPLICANT: Leiyo, Ilmo
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
; TITLE OF INVENTION: Fragments and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,077
; FILING DATE: 22-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US 94/10730
; FILING DATE: 21-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,319
; FILING DATE: 30-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/919,951
; FILING DATE: 27-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9721
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6942 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-125-077-3
;
Query Match 76.0%; Score 15.2; DB 2; Length 6942;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
Db 920 GGTCATGCCAGGCTGTGTC 939

RESULT 49
US-09-562-702A-7
; Sequence 7, Application US/09562702A
; Patent No. 6632790
; GENERAL INFORMATION:
; APPLICANT: Yurchenco, Peter
; TITLE OF INVENTION: Laminin 2 and Methods for Its Use
; FILE REFERENCE: 99-274-B
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/155,945
; PRIOR FILING DATE: 1999-09-24
; PRIOR APPLICATION NUMBER: 60/143,289
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/139,198
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/131,720
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
; LENGTH: 9419
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(9264)
; US-09-562-702A-7
;
Query Match 76.0%; Score 15.2; DB 4; Length 9419;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
Db 805 GGTCATGCCAGGCTGTGTC 824

RESULT 50
US-09-562-702A-7/c
; Sequence 7, Application US/09562702A
; Patent No. 6632790
; GENERAL INFORMATION:
; APPLICANT: Yurchenco, Peter
; TITLE OF INVENTION: Laminin 2 and Methods for Its Use
; FILE REFERENCE: 99-274-B
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/155,945
; PRIOR FILING DATE: 1999-09-24
; PRIOR APPLICATION NUMBER: 60/143,289
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/139,198
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/131,720
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-125-077-3
;
Query Match 76.0%; Score 15.2; DB 2; Length 6942;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
Db 939 GGACATGCCCTGGCATGACC 920

RESULT 49
US-09-562-702A-7
; Sequence 7, Application US/09562702A
; Patent No. 6632790
; GENERAL INFORMATION:
; APPLICANT: Yurchenco, Peter
; TITLE OF INVENTION: Laminin 2 and Methods for Its Use
; FILE REFERENCE: 99-274-B
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/155,945
; PRIOR FILING DATE: 1999-09-24
; PRIOR APPLICATION NUMBER: 60/143,289
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/139,198
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/131,720
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
;
Query Match 76.0%; Score 15.2; DB 4; Length 9419;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTCC 20
Db 805 GGTCATGCCAGGCTGTGTC 824

RESULT 50
US-09-562-702A-7/c
; Sequence 7, Application US/09562702A
; Patent No. 6632790
; GENERAL INFORMATION:
; APPLICANT: Yurchenco, Peter
; TITLE OF INVENTION: Laminin 2 and Methods for Its Use
; FILE REFERENCE: 99-274-B
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/155,945
; PRIOR FILING DATE: 1999-09-24
; PRIOR APPLICATION NUMBER: 60/143,289
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/139,198
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/131,720
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
```

;
; LENGTH: 9419
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(9264)
US-09-562-702A-7

Query Match 76.0%; Score 15.2; DB 4; Length 9419;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20
Db 824 GGACAGCCCTGGCATGACC 805

Search completed: January 14, 2005, 17:39:57
Job time : 88.0526 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 16:14:28 ; Search time 348.421 Seconds
(without alignments)
329.824 Million cell updates/sec

Title: US-09-578-453-2

Perfect score: 20
Sequence: 1 GCAGATCCCGGCGATCTCC 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 500 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCTR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	16	US-10-447-136-162 Sequence 162, App
2	20	100.0	20	16	US-10-447-136-162 Sequence 162, App
3	20	100.0	24	14	US-10-023-318-19 Sequence 19, App
4	20	100.0	24	14	US-10-023-318-19 Sequence 19, App
5	20	100.0	24	14	US-10-023-318-19 Sequence 19, App
6	20	100.0	24	14	US-10-023-318-38 Sequence 38, App
7	20	100.0	24	14	US-10-023-318-38 Sequence 38, App
8	19	95.0	19	15	US-10-339-161-14 Sequence 14, App
9	17.4	87.0	729	17	US-10-767-701-2482 Sequence 2482, Ap
10	17.4	87.0	729	17	US-10-767-701-2482 Sequence 2482, Ap
11	17.4	87.0	3325	16	US-10-425-114-30682 Sequence 30682, A
12	17.4	87.0	3325	16	US-10-425-114-30682 Sequence 30682, A

13	17.4	87.0	3458	18	US-10-425-115-4271 Sequence 4271, Ap
14	17.4	87.0	3458	18	US-10-425-115-4271 Sequence 4271, Ap
15	17.4	87.0	3569	16	US-10-425-114-2765 Sequence 2765, Ap
16	17.4	87.0	3569	16	US-10-425-114-2765 Sequence 2765, Ap
17	17.4	87.0	3609	18	US-10-425-115-4270 Sequence 4270, Ap
18	17.4	87.0	3609	18	US-10-425-115-4270 Sequence 4270, Ap
19	16.8	84.0	30	10	US-09-829-922-31 Sequence 31, App
20	16.8	84.0	30	10	US-09-829-922-31 Sequence 31, App
21	16.8	84.0	30	15	US-10-160-290-31 Sequence 31, App
22	16.8	84.0	30	15	US-10-160-290-31 Sequence 31, App
23	16.8	84.0	21565	16	US-10-052-482-61 Sequence 61, App
24	16.8	84.0	21565	16	US-10-052-482-61 Sequence 61, App
25	16.8	84.0	46030	17	US-10-367-094-1 Sequence 1, App
26	16.8	84.0	46030	17	US-10-367-094-1 Sequence 1, App
27	16.8	84.0	60327	16	US-10-052-482-187 Sequence 187, App
28	16.8	84.0	60327	16	US-10-052-482-187 Sequence 187, App
29	16.8	84.0	68233	15	US-10-034-650-31 Sequence 31, App
30	16.8	84.0	68233	15	US-10-034-650-31 Sequence 31, App
31	16.8	84.0	77941	13	US-10-087-192-709 Sequence 709, App
32	16.8	84.0	77941	13	US-10-087-192-709 Sequence 709, App
33	16.8	84.0	145068	17	US-10-322-281-33 Sequence 33, App
34	16.8	84.0	145068	17	US-10-322-281-33 Sequence 33, App
35	16.4	82.0	400	10	US-09-835-9768-77 Sequence 77, App
36	16.4	82.0	400	10	US-09-835-9768-77 Sequence 77, App
37	16.4	82.0	3422	10	US-09-835-9768-111 Sequence 111, App
38	16.4	82.0	3422	10	US-09-835-9768-111 Sequence 111, App
39	16.4	82.0	32351	17	US-10-322-696-40 Sequence 40, App
40	16.4	82.0	32351	17	US-10-322-696-40 Sequence 40, App
41	16.4	82.0	203264	13	US-10-087-192-988 Sequence 988, App
42	16.4	82.0	203264	13	US-10-087-192-988 Sequence 988, App
43	15.8	79.0	274	9	US-09-294-0938-1681 Sequence 1681, Ap
44	15.8	79.0	274	9	US-09-294-0938-1681 Sequence 1681, Ap
45	15.8	79.0	277	10	US-09-291-936-580 Sequence 580, App
46	15.8	79.0	277	10	US-09-291-936-580 Sequence 580, App
47	15.8	79.0	294	9	US-09-294-0938-5575 Sequence 5575, App
48	15.8	79.0	294	9	US-09-294-0938-5575 Sequence 5575, App
49	15.8	79.0	482	9	US-09-783-590-4831 Sequence 4831, Ap
50	15.8	79.0	482	9	US-09-783-590-4831 Sequence 4831, Ap
51	15.8	79.0	501	16	US-10-282-1228-17485 Sequence 17485, A
52	15.8	79.0	501	16	US-10-282-1228-17485 Sequence 17485, A
53	15.8	79.0	1217	16	US-10-425-114-20194 Sequence 20194, A
54	15.8	79.0	1217	16	US-10-425-114-20194 Sequence 20194, A
55	15.8	79.0	1539	15	US-10-156-761-998 Sequence 998, App
56	15.8	79.0	1539	15	US-10-156-761-998 Sequence 998, App
57	15.8	79.0	2253	18	US-10-425-115-141130 Sequence 141130, Sequence 141130,
58	15.8	79.0	2253	18	US-10-425-115-141130 Sequence 141130, Sequence 141130,
59	15.8	79.0	330973	13	US-10-087-192-1498 Sequence 1498, Ap
60	15.8	79.0	330973	13	US-10-087-192-1498 Sequence 1498, Ap
61	15.8	79.0	438892	13	US-10-087-192-454 Sequence 454, App
62	15.8	79.0	438892	13	US-10-087-192-454 Sequence 454, App
63	15.8	79.0	71651	18	US-10-719-993-6817 Sequence 6817, Ap
64	15.8	79.0	71651	18	US-10-719-993-6817 Sequence 6817, Ap
65	15.8	79.0	2731748	17	US-10-297-465A-1 Sequence 1, App
66	15.8	79.0	2731748	17	US-10-297-465A-1 Sequence 1, App
67	15.8	79.0	9025608	15	US-10-156-761-1 Sequence 1, App
68	15.8	79.0	9025608	15	US-10-156-761-1 Sequence 1, App
69	15.4	77.0	473	10	US-09-866-050A-355 Sequence 355, App
70	15.4	77.0	473	10	US-09-866-050A-355 Sequence 355, App
71	15.4	77.0	473	14	US-10-152-661-355 Sequence 355, App
72	15.4	77.0	473	14	US-10-152-661-355 Sequence 355, App
73	15.4	77.0	764	13	US-10-027-632-142160 Sequence 142160, Sequence 142160,
74	15.4	77.0	764	13	US-10-027-632-142160 Sequence 142160, Sequence 142160,
75	15.4	77.0	764	15	US-10-027-632-142160 Sequence 142160, Sequence 142160,
76	15.4	77.0	764	15	US-10-027-632-142160 Sequence 142160, Sequence 142160,
77	15.4	77.0	1701	15	US-10-369-493-39872 Sequence 39872, A
78	15.4	77.0	1701	15	US-10-369-493-39872 Sequence 39872, A
79	15.4	77.0	1857	15	US-10-369-493-39490 Sequence 39490, A
80	15.4	77.0	1857	15	US-10-369-493-39490 Sequence 39490, A
81	15.4	77.0	2315	15	US-10-369-493-39122 Sequence 39122, A
82	15.4	77.0	2315	15	US-10-369-493-39122 Sequence 39122, A
83	15.4	77.0	7282	15	US-10-213-948-11 Sequence 11, App
84	15.4	77.0	7282	15	US-10-213-948-11 Sequence 11, App
85	15.4	77.0	57013	13	US-10-087-192-1798 Sequence 1798, Ap

C 86	15.4	77.0	57013	13	US-10-087-132-1798	Sequence 1798, Ap	159	15.2	76.0	3078	15	US-10-027-632-177166	Sequence 177166, Ap
C 87	15.4	77.0	114411	18	US-10-723-860-3910	Sequence 3910, Ap	C 160	15.2	76.0	3078	15	US-10-027-632-177166	Sequence 177166, Ap
C 88	15.4	77.0	114411	18	US-10-723-860-3910	Sequence 3910, Ap	C 161	15.2	76.0	4258	15	US-10-128-714-494	Sequence 494, Ap
C 89	15.2	76.0	297	9	US-09-878-722-72	Sequence 72, Appl	C 162	15.2	76.0	4258	15	US-10-128-714-494	Sequence 494, Ap
C 90	15.2	76.0	297	9	US-09-878-722-72	Sequence 72, Appl	C 163	15.2	76.0	4258	15	US-10-128-714-494	Sequence 494, Ap
C 91	15.2	76.0	297	9	US-09-878-178-1454	Sequence 1454, Ap	C 164	15.2	76.0	4258	15	US-10-128-714-494	Sequence 494, Ap
C 92	15.2	76.0	297	9	US-09-878-178-1454	Sequence 1454, Ap	C 165	15.2	76.0	5340	17	US-10-437-963-9004	Sequence 9004, Ap
C 93	15.2	76.0	297	10	US-09-904-456-72	Sequence 72, Appl	C 166	15.2	76.0	5340	17	US-10-437-963-9004	Sequence 9004, Ap
C 94	15.2	76.0	297	10	US-09-904-456-72	Sequence 72, Appl	C 167	15.2	76.0	9534	9	US-09-954-531-1380	Sequence 1380, Ap
C 95	15.2	76.0	297	13	US-10-046-935-1454	Sequence 1454, Ap	C 168	15.2	76.0	9534	9	US-09-954-531-1380	Sequence 1380, Ap
C 96	15.2	76.0	297	13	US-10-046-935-1454	Sequence 1454, Ap	C 169	15.2	76.0	9649	18	US-10-723-860-697	Sequence 697, App
C 97	15.2	76.0	297	14	US-10-146-502-1454	Sequence 1454, Ap	C 170	15.2	76.0	9649	18	US-10-723-860-697	Sequence 697, App
C 98	15.2	76.0	297	14	US-10-146-502-1454	Sequence 1454, Ap	C 171	15.2	76.0	9670	18	US-10-723-860-697	Sequence 697, App
C 99	15.2	76.0	298	18	US-10-425-115-39745	Sequence 39745, A	C 172	15.2	76.0	9670	18	US-10-723-860-697	Sequence 697, App
C 100	15.2	76.0	298	18	US-10-425-115-39745	Sequence 39745, A	C 173	15.2	76.0	10867	9	US-09-764-869-2113	Sequence 2113, Ap
C 101	15.2	76.0	366	18	US-10-744-936-4	Sequence 4, Appl1	C 174	15.2	76.0	10867	9	US-09-764-869-2113	Sequence 2113, Ap
C 102	15.2	76.0	366	18	US-10-744-936-4	Sequence 4, Appl1	C 175	15.2	76.0	10867	9	US-09-764-869-2113	Sequence 2113, Ap
C 103	15.2	76.0	401	10	US-09-918-995-4352	Sequence 4352, Ap	C 176	15.2	76.0	10867	9	US-09-764-869-2113	Sequence 2113, Ap
C 104	15.2	76.0	401	10	US-09-918-995-4352	Sequence 4352, Ap	C 177	15.2	76.0	10867	14	US-10-091-504-2113	Sequence 2113, Ap
C 105	15.2	76.0	456	16	US-10-424-539-60443	Sequence 60443, A	C 178	15.2	76.0	10867	14	US-10-091-504-2113	Sequence 2113, Ap
C 106	15.2	76.0	456	16	US-10-424-539-60443	Sequence 60443, A	C 179	15.2	76.0	10867	16	US-10-227-577-2113	Sequence 2113, Ap
C 107	15.2	76.0	472	10	US-09-918-995-13145	Sequence 13145, A	C 180	15.2	76.0	10867	16	US-10-227-577-2113	Sequence 2113, Ap
C 108	15.2	76.0	472	10	US-09-918-995-13145	Sequence 13145, A	C 181	15.2	76.0	10867	16	US-10-242-515-3970	Sequence 3970, Ap
C 109	15.2	76.0	491	10	US-09-918-995-11130	Sequence 11130, A	C 182	15.2	76.0	10867	16	US-10-242-515-3970	Sequence 3970, Ap
C 110	15.2	76.0	491	10	US-09-918-995-11130	Sequence 11130, A	C 183	15.2	76.0	10894	9	US-09-764-869-2111	Sequence 2111, Ap
C 111	15.2	76.0	492	10	US-09-918-995-1921	Sequence 1921, Ap	C 184	15.2	76.0	10894	9	US-09-764-869-2111	Sequence 2111, Ap
C 112	15.2	76.0	492	10	US-09-918-995-1921	Sequence 1921, Ap	C 185	15.2	76.0	10894	14	US-10-091-504-2111	Sequence 2111, Ap
C 113	15.2	76.0	498	15	US-10-156-761-1173	Sequence 1173, Ap	C 186	15.2	76.0	10894	14	US-10-091-504-2111	Sequence 2111, Ap
C 114	15.2	76.0	498	15	US-10-156-761-1173	Sequence 1173, Ap	C 187	15.2	76.0	10894	16	US-10-227-577-2111	Sequence 2111, Ap
C 115	15.2	76.0	516	16	US-10-424-539-106147	Sequence 106147, A	C 188	15.2	76.0	10894	16	US-10-227-577-2111	Sequence 2111, Ap
C 116	15.2	76.0	516	16	US-10-424-539-106147	Sequence 106147, A	C 189	15.2	76.0	10907	9	US-09-764-869-2112	Sequence 2112, Ap
C 117	15.2	76.0	538	15	US-10-029-386-13284	Sequence 13284, A	C 190	15.2	76.0	10907	9	US-09-764-869-2112	Sequence 2112, Ap
C 118	15.2	76.0	538	15	US-10-029-386-13284	Sequence 13284, A	C 191	15.2	76.0	10907	9	US-09-764-869-2112	Sequence 2112, Ap
C 119	15.2	76.0	569	15	US-10-029-386-13639	Sequence 13639, A	C 192	15.2	76.0	10907	9	US-09-764-869-2112	Sequence 2112, Ap
C 120	15.2	76.0	569	15	US-10-029-386-13639	Sequence 13639, A	C 193	15.2	76.0	10907	14	US-10-091-504-2112	Sequence 2112, Ap
C 121	15.2	76.0	574	15	US-10-029-386-7058	Sequence 7058, Ap	C 194	15.2	76.0	10907	14	US-10-091-504-2112	Sequence 2112, Ap
C 122	15.2	76.0	574	15	US-10-029-386-7058	Sequence 7058, Ap	C 195	15.2	76.0	10907	16	US-10-227-577-2112	Sequence 2112, Ap
C 123	15.2	76.0	586	18	US-10-425-115-78128	Sequence 78128, A	C 196	15.2	76.0	10907	16	US-10-227-577-2112	Sequence 2112, Ap
C 124	15.2	76.0	586	18	US-10-425-115-78128	Sequence 78128, A	C 197	15.2	76.0	10907	16	US-10-242-515-3968	Sequence 3968, Ap
C 125	15.2	76.0	615	15	US-10-163-198-32	Sequence 32, Appl	C 198	15.2	76.0	10907	16	US-10-242-515-3968	Sequence 3968, Ap
C 126	15.2	76.0	615	15	US-10-163-198-32	Sequence 32, Appl	C 199	15.2	76.0	12718	9	US-09-764-877-3972	Sequence 3972, Ap
C 127	15.2	76.0	627	16	US-10-424-539-104831	Sequence 104831, A	C 200	15.2	76.0	12718	9	US-09-764-877-3972	Sequence 3972, Ap
C 128	15.2	76.0	627	16	US-10-424-539-104831	Sequence 104831, A	C 201	15.2	76.0	12718	16	US-10-242-515-3972	Sequence 3972, Ap
C 129	15.2	76.0	1008	15	US-10-156-761-2985	Sequence 2985, Ap	C 202	15.2	76.0	12718	16	US-10-242-515-3972	Sequence 3972, Ap
C 130	15.2	76.0	1008	15	US-10-156-761-2985	Sequence 2985, Ap	C 203	15.2	76.0	33097	13	US-10-087-199-241	Sequence 241, App
C 131	15.2	76.0	1152	15	US-10-259-165-201	Sequence 201, App	C 204	15.2	76.0	33097	13	US-10-087-199-241	Sequence 241, App
C 132	15.2	76.0	1152	15	US-10-259-165-201	Sequence 201, App	C 205	15.2	76.0	39405	13	US-10-087-199-241	Sequence 241, App
C 133	15.2	76.0	1251	10	US-09-828-000-1	Sequence 1, Appl1	C 206	15.2	76.0	39405	13	US-10-087-199-241	Sequence 241, App
C 134	15.2	76.0	1251	10	US-09-828-000-1	Sequence 1, Appl1	C 207	15.2	76.0	50720	17	US-10-322-281-621	Sequence 621, App
C 135	15.2	76.0	1251	15	US-10-405-588-1	Sequence 1, Appl1	C 208	15.2	76.0	50720	17	US-10-322-281-621	Sequence 621, App
C 136	15.2	76.0	1251	15	US-10-405-588-1	Sequence 1, Appl1	C 209	15.2	76.0	67191	11	US-09-997-722-169	Sequence 169, App
C 137	15.2	76.0	1308	15	US-10-156-761-752	Sequence 752, App	C 210	15.2	76.0	67191	11	US-09-997-722-169	Sequence 169, App
C 138	15.2	76.0	1308	15	US-10-156-761-752	Sequence 752, App	C 211	15.2	76.0	67191	15	US-10-105-612-1	Sequence 1, Appl1
C 139	15.2	76.0	1513	16	US-10-424-539-104834	Sequence 104834, A	C 212	15.2	76.0	67191	15	US-10-105-612-1	Sequence 1, Appl1
C 140	15.2	76.0	1513	16	US-10-424-539-104834	Sequence 104834, A	C 213	15.2	76.0	110096	9	US-09-880-107-1542	Sequence 1542, Ap
C 141	15.2	76.0	1569	17	US-10-437-963-53491	Sequence 53491, A	C 214	15.2	76.0	110096	9	US-09-880-107-1542	Sequence 1542, Ap
C 142	15.2	76.0	1569	17	US-10-437-963-53491	Sequence 53491, A	C 215	15.2	76.0	163701	17	US-10-322-281-439	Sequence 439, App
C 143	15.2	76.0	1716	13	US-10-027-632-249965	Sequence 249965, A	C 216	15.2	76.0	163701	17	US-10-322-281-439	Sequence 439, App
C 144	15.2	76.0	1716	13	US-10-027-632-249965	Sequence 249965, A	C 217	15.2	76.0	173808	13	US-10-003-806-10	Sequence 10, Appl
C 145	15.2	76.0	1716	15	US-10-027-632-249965	Sequence 249965, A	C 218	15.2	76.0	173808	13	US-10-003-806-10	Sequence 10, Appl
C 146	15.2	76.0	1716	15	US-10-027-632-249965	Sequence 249965, A	C 219	15.2	76.0	347001	17	US-10-013-908-16	Sequence 16, Appl
C 147	15.2	76.0	1899	16	US-10-439-703-7	Sequence 7, Appl1	C 220	15.2	76.0	347001	17	US-10-013-908-16	Sequence 16, Appl
C 148	15.2	76.0	1899	16	US-10-439-703-7	Sequence 7, Appl1	C 221	14.8	74.0	49	9	US-09-813-824-31	Sequence 31, Appl
C 149	15.2	76.0	1899	16	US-09-906-393A-35	Sequence 35, Appl	C 222	14.8	74.0	49	9	US-09-813-824-31	Sequence 31, Appl
C 150	15.2	76.0	1920	10	US-09-906-393A-35	Sequence 35, Appl	C 223	14.8	74.0	201	17	US-10-741-601-1449	Sequence 14449, A
C 151	15.2	76.0	1924	15	US-10-240-965-465	Sequence 265, App	C 224	14.8	74.0	201	17	US-10-741-601-1449	Sequence 14449, A
C 152	15.2	76.0	1924	15	US-10-240-965-465	Sequence 265, App	C 225	14.8	74.0	400	16	US-10-424-539-90906	Sequence 90906, A
C 153	15.2	76.0	1940	15	US-10-367-093-13	Sequence 13, Appl	C 226	14.8	74.0	400	16	US-10-424-539-90906	Sequence 90906, A
C 154	15.2	76.0	1940	15	US-10-367-093-13	Sequence 13, Appl	C 227	14.8	74.0	478	10	US-09-918-995-23256	Sequence 23256, A
C 155	15.2	76.0	1958	9	US-09-791-406-3	Sequence 3, Appl1	C 228	14.8	74.0	478	10	US-09-918-995-23256	Sequence 23256, A
C 156	15.2	76.0	1958	9	US-09-791-406-3	Sequence 3, Appl1	C 229	14.8	74.0	566	9	US-09-920-908-647	Sequence 647, App
C 157	15.2	76.0	3078	13	US-10-027-632-177166	Sequence 177166, Ap	C 230	14.8	74.0	566	9	US-09-920-908-647	Sequence 647, App
C 158	15.2	76.0	3078	13	US-10-027-632-177166	Sequence 177166, Ap	C 231	14.8	74.0	566	13	US-10-033-500-647	Sequence 647, App

C 232	14.8	74.0	566	13	US-10-033-528-647	Sequence 647, App	305	14.4	72.0	792	13	US-10-027-632-14097	Sequence 14097, A
C 233	14.8	74.0	566	15	US-10-099-926-647	Sequence 647, App	306	14.4	72.0	792	13	US-10-027-632-14097	Sequence 14097, A
C 234	14.8	74.0	566	15	US-10-099-926-647	Sequence 647, App	307	14.4	72.0	792	13	US-10-027-632-14098	Sequence 14098, A
C 235	14.8	74.0	747	9	US-09-813-453A-35	Sequence 35, App1	308	14.4	72.0	792	13	US-10-027-632-14098	Sequence 14098, A
C 236	14.8	74.0	747	9	US-09-813-453A-35	Sequence 35, App1	309	14.4	72.0	792	13	US-10-027-632-14099	Sequence 14099, A
C 237	14.8	74.0	789	13	US-10-027-632-12458	Sequence 12458, A	310	14.4	72.0	792	13	US-10-027-632-14099	Sequence 14099, A
C 238	14.8	74.0	789	13	US-10-027-632-12458	Sequence 12458, A	311	14.4	72.0	792	15	US-10-027-632-14096	Sequence 14096, A
C 239	14.8	74.0	789	15	US-10-027-632-12458	Sequence 12458, A	312	14.4	72.0	792	15	US-10-027-632-14096	Sequence 14096, A
C 240	14.8	74.0	789	15	US-10-027-632-12458	Sequence 12458, A	313	14.4	72.0	792	15	US-10-027-632-14097	Sequence 14097, A
C 241	14.8	74.0	811	13	US-10-159-749-5	Sequence 5, App1	314	14.4	72.0	792	15	US-10-027-632-14097	Sequence 14097, A
C 242	14.8	74.0	811	13	US-10-159-749-5	Sequence 5, App1	315	14.4	72.0	792	15	US-10-027-632-14098	Sequence 14098, A
C 243	14.8	74.0	994	18	US-10-653-047-30	Sequence 30, App1	316	14.4	72.0	792	15	US-10-027-632-14098	Sequence 14098, A
C 244	14.8	74.0	994	18	US-10-653-047-30	Sequence 30, App1	317	14.4	72.0	792	15	US-10-027-632-14099	Sequence 14099, A
C 245	14.8	74.0	1029	18	US-10-425-115-95380	Sequence 95380, A	318	14.4	72.0	792	15	US-10-027-632-14099	Sequence 14099, A
C 246	14.8	74.0	1029	18	US-10-425-115-95380	Sequence 95380, A	319	14.4	72.0	1150	16	US-10-260-238-3780	Sequence 3780, App
C 247	14.8	74.0	1057	18	US-10-425-115-14357	Sequence 14357, A	320	14.4	72.0	1150	16	US-10-260-238-3780	Sequence 3780, App
C 248	14.8	74.0	1057	18	US-10-425-115-14357	Sequence 14357, A	321	14.4	72.0	1611	15	US-10-369-493-44284	Sequence 44284, A
C 249	14.8	74.0	1354	17	US-10-437-963-71006	Sequence 71006, A	322	14.4	72.0	1611	15	US-10-369-493-44284	Sequence 44284, A
C 250	14.8	74.0	1354	17	US-10-437-963-71006	Sequence 71006, A	323	14.4	72.0	1651	17	US-10-767-701-13956	Sequence 13956, A
C 251	14.8	74.0	1725	17	US-10-437-963-60963	Sequence 60963, A	324	14.4	72.0	1651	17	US-10-767-701-13956	Sequence 13956, A
C 252	14.8	74.0	1725	17	US-10-437-963-60963	Sequence 60963, A	325	14.4	72.0	1677	15	US-10-369-493-33181	Sequence 33181, A
C 253	14.8	74.0	1781	15	US-10-316-754-20	Sequence 20, App1	326	14.4	72.0	1677	15	US-10-369-493-33181	Sequence 33181, A
C 254	14.8	74.0	1781	15	US-10-316-754-20	Sequence 20, App1	327	14.4	72.0	2002	14	US-10-153-668-173	Sequence 173, App
C 255	14.8	74.0	1949	15	US-10-316-754-12	Sequence 12, App1	328	14.4	72.0	2002	14	US-10-153-668-173	Sequence 173, App
C 256	14.8	74.0	1949	15	US-10-316-754-12	Sequence 12, App1	329	14.4	72.0	3744	14	US-10-153-668-173	Sequence 465, App
C 257	14.8	74.0	2044	16	US-10-425-114-31320	Sequence 31320, A	330	14.4	72.0	3744	14	US-10-153-668-173	Sequence 465, App
C 258	14.8	74.0	2044	16	US-10-425-114-31320	Sequence 31320, A	331	14.4	72.0	3901	14	US-10-153-668-173	Sequence 175, App
C 259	14.8	74.0	2087	16	US-10-425-114-4515	Sequence 4515, App	332	14.4	72.0	3901	14	US-10-153-668-173	Sequence 175, App
C 260	14.8	74.0	2087	16	US-10-425-114-4515	Sequence 4515, App	333	14.4	72.0	5066	10	US-09-824-574-1	Sequence 1, App1
C 261	14.8	74.0	2088	15	US-10-156-761-1788	Sequence 1788, App	334	14.4	72.0	5066	10	US-09-824-574-1	Sequence 1, App1
C 262	14.8	74.0	2088	15	US-10-156-761-1788	Sequence 1788, App	335	14.4	72.0	30000	12	US-09-980-217-2	Sequence 2, App1
C 263	14.8	74.0	2152	18	US-10-425-115-95382	Sequence 95382, A	336	14.4	72.0	30000	12	US-09-980-217-2	Sequence 2, App1
C 264	14.8	74.0	2152	18	US-10-425-115-95382	Sequence 95382, A	337	14.4	72.0	40394	17	US-10-741-601-5774	Sequence 5774, App
C 265	14.8	74.0	3621	17	US-10-437-963-12579	Sequence 12579, A	338	14.4	72.0	40394	17	US-10-741-601-5774	Sequence 5774, App
C 266	14.8	74.0	3621	17	US-10-437-963-12579	Sequence 12579, A	339	14.4	72.0	59944	17	US-10-741-601-5619	Sequence 5619, App
C 267	14.8	74.0	3934	9	US-09-813-453A-77	Sequence 77, App1	340	14.4	72.0	59944	17	US-10-741-601-5619	Sequence 5619, App
C 268	14.8	74.0	3934	9	US-09-813-453A-77	Sequence 77, App1	341	14.4	72.0	101782	17	US-10-741-601-5611	Sequence 5611, App
C 269	14.8	74.0	63693	17	US-10-741-601-5650	Sequence 5650, App	342	14.4	72.0	101782	17	US-10-741-601-5611	Sequence 5661, App
C 270	14.8	74.0	63693	17	US-10-741-601-5650	Sequence 5650, App	343	14.4	72.0	157044	13	US-10-087-192-565	Sequence 565, App
C 271	14.8	74.0	87350	16	US-10-374-077-79	Sequence 79, App1	344	14.4	72.0	157044	13	US-10-087-192-565	Sequence 565, App
C 272	14.8	74.0	87350	16	US-10-374-077-79	Sequence 79, App1	345	14.4	72.0	2256646	17	US-10-470-565-1	Sequence 1, App1
C 273	14.8	74.0	151870	17	US-10-741-601-5614	Sequence 5614, App	346	14.4	72.0	2256646	17	US-10-470-565-1	Sequence 1, App1
C 274	14.8	74.0	151870	17	US-10-741-601-5614	Sequence 5614, App	347	14.2	71.0	183	17	US-10-437-963-100008	Sequence 100008, A
C 275	14.8	74.0	313287	17	US-10-322-281-48	Sequence 48, App1	348	14.2	71.0	183	17	US-10-437-963-100008	Sequence 100008, A
C 276	14.8	74.0	313287	17	US-10-322-281-48	Sequence 48, App1	349	14.2	71.0	201	18	US-10-719-993-41925	Sequence 41925, A
C 277	14.8	74.0	1601042	13	US-10-027-632-55064	Sequence 55064, A	350	14.2	71.0	201	18	US-10-719-993-41925	Sequence 41925, A
C 278	14.8	74.0	1601042	13	US-10-027-632-55064	Sequence 55064, A	351	14.2	71.0	228	18	US-10-425-115-116704	Sequence 116704, A
C 279	14.8	74.0	1601042	15	US-10-027-632-55064	Sequence 55064, A	352	14.2	71.0	228	18	US-10-425-115-116704	Sequence 116704, A
C 280	14.8	74.0	1601042	15	US-10-027-632-55064	Sequence 55064, A	353	14.2	71.0	237	17	US-10-469-285-383	Sequence 383, App
C 281	14.4	72.0	187	18	US-10-653-047-7853	Sequence 7853, App	354	14.2	71.0	237	17	US-10-469-285-383	Sequence 383, App
C 282	14.4	72.0	187	18	US-10-653-047-7853	Sequence 7853, App	355	14.2	71.0	250	11	US-09-864-408A-4661	Sequence 4661, App
C 283	14.4	72.0	200	18	US-10-425-115-159810	Sequence 159810, A	356	14.2	71.0	250	11	US-09-864-408A-4661	Sequence 4661, App
C 284	14.4	72.0	200	18	US-10-425-115-159810	Sequence 159810, A	357	14.2	71.0	261	15	US-10-245-618-9	Sequence 9, App1
C 285	14.4	72.0	352	18	US-10-425-115-87386	Sequence 87386, A	358	14.2	71.0	261	15	US-10-245-618-9	Sequence 9, App1
C 286	14.4	72.0	352	18	US-10-425-115-87386	Sequence 87386, A	359	14.2	71.0	261	15	US-10-245-618-9	Sequence 29, App1
C 287	14.4	72.0	370	15	US-10-106-698-2499	Sequence 2499, App	360	14.2	71.0	261	15	US-10-245-618-9	Sequence 29, App1
C 288	14.4	72.0	370	15	US-10-106-698-2499	Sequence 2499, App	361	14.2	71.0	294	18	US-10-425-115-154546	Sequence 154546, A
C 289	14.4	72.0	399	18	US-10-425-115-130668	Sequence 130668, A	362	14.2	71.0	294	18	US-10-425-115-154546	Sequence 154546, A
C 290	14.4	72.0	399	18	US-10-425-115-130668	Sequence 130668, A	363	14.2	71.0	360	16	US-10-282-122A-19545	Sequence 19545, A
C 291	14.4	72.0	453	18	US-10-425-115-128493	Sequence 128493, A	364	14.2	71.0	360	16	US-10-282-122A-19545	Sequence 19545, A
C 292	14.4	72.0	453	18	US-10-425-115-128493	Sequence 128493, A	365	14.2	71.0	427	18	US-10-425-115-173311	Sequence 173311, A
C 293	14.4	72.0	473	16	US-10-260-238-3779	Sequence 3779, App	366	14.2	71.0	427	18	US-10-425-115-173311	Sequence 173311, A
C 294	14.4	72.0	473	16	US-10-260-238-3779	Sequence 3779, App	367	14.2	71.0	393	18	US-10-425-115-130810	Sequence 130810, A
C 295	14.4	72.0	486	17	US-10-767-701-22567	Sequence 22567, App	368	14.2	71.0	393	18	US-10-425-115-130810	Sequence 130810, A
C 296	14.4	72.0	486	17	US-10-767-701-22567	Sequence 22567, App	369	14.2	71.0	419	18	US-10-425-115-127592	Sequence 127592, A
C 297	14.4	72.0	697	16	US-10-260-238-3781	Sequence 3781, App	370	14.2	71.0	419	18	US-10-425-115-127592	Sequence 127592, A
C 298	14.4	72.0	697	16	US-10-260-238-3781	Sequence 3781, App	371	14.2	71.0	427	18	US-10-425-115-173311	Sequence 173311, A
C 299	14.4	72.0	725	13	US-10-027-632-99538	Sequence 99538, A	372	14.2	71.0	427	18	US-10-425-115-173311	Sequence 173311, A
C 300	14.4	72.0	725	13	US-10-027-632-99538	Sequence 99538, A	373	14.2	71.0	466	9	US-09-738-626-941	Sequence 941, App
C 301	14.4	72.0	725	13	US-10-027-632-99538	Sequence 99538, A	374	14.2	71.0	466	9	US-09-738-626-941	Sequence 941, App
C 302	14.4	72.0	725	15	US-10-027-632-99538	Sequence 99538, A	375	14.2	71.0	466	9	US-09-738-626-941	Sequence 941, App
C 303	14.4	72.0	792	13	US-10-027-632-14096	Sequence 14096, A	376	14.2	71.0	459	18	US-10-425-115-55280	Sequence 55280, A
C 304	14.4	72.0	792	13	US-10-027-632-14096	Sequence 14096, A	377	14.2	71.0	464	18	US-10-425-115-82415	Sequence 82415, A

C 378	14.2	71.0	464	18	US-10-425-115-82415	Sequence 82415, A
C 379	14.2	71.0	476	13	US-10-027-632-50228	Sequence 50228, A
C 380	14.2	71.0	476	13	US-10-027-632-50228	Sequence 50228, A
C 381	14.2	71.0	476	13	US-10-027-632-50229	Sequence 50229, A
C 382	14.2	71.0	476	13	US-10-027-632-50229	Sequence 50229, A
C 383	14.2	71.0	476	13	US-10-027-632-50230	Sequence 50230, A
C 384	14.2	71.0	476	13	US-10-027-632-50230	Sequence 50230, A
C 385	14.2	71.0	476	13	US-10-027-632-50228	Sequence 50228, A
C 386	14.2	71.0	476	13	US-10-027-632-50228	Sequence 50228, A
C 387	14.2	71.0	476	13	US-10-027-632-50229	Sequence 50229, A
C 388	14.2	71.0	476	13	US-10-027-632-50229	Sequence 50229, A
C 389	14.2	71.0	476	13	US-10-027-632-50230	Sequence 50230, A
C 390	14.2	71.0	476	13	US-10-027-632-50230	Sequence 50230, A
C 391	14.2	71.0	485	18	US-10-425-115-161855	Sequence 161855, A
C 392	14.2	71.0	485	18	US-10-425-115-161855	Sequence 161855, A
C 393	14.2	71.0	531	16	US-10-425-114-7176	Sequence 7176, A
C 394	14.2	71.0	531	16	US-10-425-114-7176	Sequence 7176, A
C 395	14.2	71.0	531	16	US-10-425-115-108651	Sequence 108651, A
C 396	14.2	71.0	531	16	US-10-425-115-108651	Sequence 108651, A
C 397	14.2	71.0	563	16	US-10-260-238-3142	Sequence 3142, A
C 398	14.2	71.0	563	16	US-10-260-238-3142	Sequence 3142, A
C 399	14.2	71.0	580	16	US-10-260-238-4658	Sequence 4658, A
C 400	14.2	71.0	580	16	US-10-260-238-4658	Sequence 4658, A
C 401	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 402	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 403	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 404	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 405	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 406	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 407	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 408	14.2	71.0	583	13	US-10-027-632-113405	Sequence 113405, A
C 409	14.2	71.0	589	16	US-10-260-238-3130	Sequence 3130, A
C 410	14.2	71.0	589	16	US-10-260-238-3130	Sequence 3130, A
C 411	14.2	71.0	590	18	US-10-357-930-59344	Sequence 59344, A
C 412	14.2	71.0	590	18	US-10-357-930-59344	Sequence 59344, A
C 413	14.2	71.0	591	15	US-10-006-285-309	Sequence 309, A
C 414	14.2	71.0	591	15	US-10-006-285-309	Sequence 309, A
C 415	14.2	71.0	598	13	US-10-027-632-190503	Sequence 190503, A
C 416	14.2	71.0	598	13	US-10-027-632-190503	Sequence 190503, A
C 417	14.2	71.0	598	13	US-10-027-632-204130	Sequence 204130, A
C 418	14.2	71.0	598	13	US-10-027-632-204130	Sequence 204130, A
C 419	14.2	71.0	598	13	US-10-027-632-190503	Sequence 190503, A
C 420	14.2	71.0	598	13	US-10-027-632-190503	Sequence 190503, A
C 421	14.2	71.0	598	13	US-10-027-632-204130	Sequence 204130, A
C 422	14.2	71.0	598	13	US-10-027-632-204130	Sequence 204130, A
C 423	14.2	71.0	600	15	US-10-029-386-2368	Sequence 2368, A
C 424	14.2	71.0	600	15	US-10-029-386-2368	Sequence 2368, A
C 425	14.2	71.0	611	13	US-10-027-632-195852	Sequence 195852, A
C 426	14.2	71.0	611	13	US-10-027-632-195852	Sequence 195852, A
C 427	14.2	71.0	611	13	US-10-027-632-195852	Sequence 195852, A
C 428	14.2	71.0	611	13	US-10-027-632-195852	Sequence 195852, A
C 429	14.2	71.0	611	13	US-10-027-632-82854	Sequence 82854, A
C 430	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 431	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 432	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 433	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 434	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 435	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 436	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 437	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 438	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 439	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 440	14.2	71.0	633	13	US-10-027-632-82854	Sequence 82854, A
C 441	14.2	71.0	636	16	US-10-260-238-3129	Sequence 3129, A
C 442	14.2	71.0	636	16	US-10-260-238-3129	Sequence 3129, A
C 443	14.2	71.0	636	16	US-10-653-047-6175	Sequence 6175, A
C 444	14.2	71.0	636	16	US-10-653-047-6175	Sequence 6175, A
C 445	14.2	71.0	651	13	US-10-027-632-99328	Sequence 99328, A
C 446	14.2	71.0	651	13	US-10-027-632-99328	Sequence 99328, A
C 447	14.2	71.0	651	13	US-10-027-632-99328	Sequence 99328, A
C 448	14.2	71.0	651	13	US-10-027-632-99328	Sequence 99328, A
C 449	14.2	71.0	657	15	US-10-282-122A-25895	Sequence 25895, A
C 450	14.2	71.0	657	15	US-10-282-122A-25895	Sequence 25895, A

C 451	14.2	71.0	697	13	US-10-001-843-108	Sequence 108, A
C 452	14.2	71.0	697	13	US-10-001-843-108	Sequence 108, A
C 453	14.2	71.0	702	18	US-10-425-115-148339	Sequence 148339, A
C 454	14.2	71.0	702	18	US-10-425-115-148339	Sequence 148339, A
C 455	14.2	71.0	724	18	US-10-425-115-145860	Sequence 145860, A
C 456	14.2	71.0	724	18	US-10-425-115-145860	Sequence 145860, A
C 457	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 458	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 459	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 460	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 461	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 462	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 463	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 464	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 465	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 466	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 467	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 468	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 469	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 470	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 471	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 472	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 473	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 474	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 475	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 476	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 477	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 478	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 479	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 480	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 481	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 482	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 483	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 484	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 485	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 486	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 487	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 488	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 489	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 490	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 491	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 492	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 493	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 494	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 495	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 496	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 497	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 498	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 499	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A
C 500	14.2	71.0	728	13	US-10-027-632-148969	Sequence 148969, A

RESULT 1
 US-10-447-136-162
 ; Sequence 162, Application US/10447136
 ; Publication No. US2004000948A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WRIGHT, Jim A.
 ; TITLE OF INVENTION: Antitumor Antisense Sequences Directed Against R1 and
 ; TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase
 ; FILE REFERENCE: 032396-023
 ; CURRENT APPLICATION NUMBER: US/10/447,136
 ; PRIOR FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/249,247
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-11
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/023,040
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-02
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/039,959
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1997-03-07

;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 08/904,901
;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-01
;; NUMBER OF SEQ ID NOS: 220
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 162
;; LENGTH: 20
;; TYPE: DNA
;; ORGANISM: Human
US-10-447-136-162

Query Match 100.0%; Score 20; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20
DB 1 GGACATGCCGGGCGCATGTCC 20

RESULT 2

US-10-447-136-162/c
;; Sequence 162, Application US/10447136
;; Publication No. US2004009948A1
;; GENERAL INFORMATION:
;; APPLICANT: WRIGHT, Jim A.
;; APPLICANT: YOUNG, Alping H.
;; TITLE OF INVENTION: Anticancer Antisense Sequences Directed Against R1 and
;; TITLE OF INVENTION: R2 Components of Ribonucleotide Reductase
;; FILE REFERENCE: 032396-023
;; CURRENT APPLICATION NUMBER: US/10/447,136
;; PRIOR FILING DATE: 2003-05-29
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/249,247
;; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-11
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/023,040
;; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-02
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/039,959
;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-03-07
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 08/904,901
;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-01
;; NUMBER OF SEQ ID NOS: 220
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 162
;; LENGTH: 20
;; TYPE: DNA
;; ORGANISM: Human
US-10-447-136-162

Query Match 100.0%; Score 20; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20
DB 20 GGACATGCCGGGCGCATGTCC 1

RESULT 3

US-10-023-318-19
;; Sequence 19, Application US/10023318
;; Publication No. US20030092015A1
;; GENERAL INFORMATION:
;; APPLICANT: Larose, Anne-Marie
;; APPLICANT: Rousseau, Pierre
;; APPLICANT: Leblanc, Benoit
;; APPLICANT: Camato, Rino
;; TITLE OF INVENTION: Method for Screening and/or Identifying Factors that
;; TITLE OF INVENTION: Bind to Nucleic Acids
;; FILE REFERENCE: 9555.123USU1
;; CURRENT APPLICATION NUMBER: US/10/023,318
;; PRIOR FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: 2,327,561
;; PRIOR FILING DATE: 2000-12-27
;; NUMBER OF SEQ ID NOS: 41

;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 19
;; LENGTH: 24
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: NABE-probes
US-10-023-318-19

Query Match 100.0%; Score 20; DB 14; Length 24;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20
DB 5 GGACATGCCGGGCGCATGTCC 24

RESULT 4

US-10-023-318-19/c
;; Sequence 19, Application US/10023318
;; Publication No. US20030092015A1
;; GENERAL INFORMATION:
;; APPLICANT: Larose, Anne-Marie
;; APPLICANT: Rousseau, Pierre
;; APPLICANT: Leblanc, Benoit
;; APPLICANT: Camato, Rino
;; TITLE OF INVENTION: Method for Screening and/or Identifying Factors that
;; TITLE OF INVENTION: Bind to Nucleic Acids
;; FILE REFERENCE: 9555.123USU1
;; CURRENT APPLICATION NUMBER: US/10/023,318
;; PRIOR FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: 2,327,561
;; PRIOR FILING DATE: 2000-12-27
;; NUMBER OF SEQ ID NOS: 41
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 19
;; LENGTH: 24
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: NABE-probes
US-10-023-318-19

Query Match 100.0%; Score 20; DB 14; Length 24;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACATGCCGGGCGCATGTCC 20
DB 24 GGACATGCCGGGCGCATGTCC 5

RESULT 5

US-10-023-318-38
;; Sequence 38, Application US/10023318
;; Publication No. US20030092015A1
;; GENERAL INFORMATION:
;; APPLICANT: Larose, Anne-Marie
;; APPLICANT: Rousseau, Pierre
;; APPLICANT: Leblanc, Benoit
;; APPLICANT: Camato, Rino
;; TITLE OF INVENTION: Method for Screening and/or Identifying Factors that
;; TITLE OF INVENTION: Bind to Nucleic Acids
;; FILE REFERENCE: 9555.123USU1
;; CURRENT APPLICATION NUMBER: US/10/023,318
;; PRIOR FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: 2,327,561
;; PRIOR FILING DATE: 2000-12-27
;; NUMBER OF SEQ ID NOS: 41
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 38
;; LENGTH: 24

```
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Double
US-10-023-318-38
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 14; Length 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGACATGCCCGGGCATGTCC 20
DB 5 GGACATGCCCGGGCATGTCC 24
```

```
RESULT 6
US-10-023-318-38/c
Sequence 38, Application US/10023318
Publication No. US20030092015A1
GENERAL INFORMATION:
APPLICANT: Larose, Anne-Marie
APPLICANT: Rousseau, Pierre
APPLICANT: Leblanc, Benoit
TITLE OF INVENTION: Method for Screening and/or Identifying Factors that
FILE REFERENCE: 9555.123USU1
CURRENT APPLICATION NUMBER: US/10/023,318
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: 2,327,561
PRIOR FILING DATE: 2000-12-27
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 38
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Double
US-10-023-318-38
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 14; Length 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGACATGCCCGGGCATGTCC 20
DB 24 GGACATGCCCGGGCATGTCC 5
```

```
RESULT 7
US-10-339-161-14
Sequence 14, Application US/10339161
Publication No. US20030162211A1
GENERAL INFORMATION:
APPLICANT: Remacle, Jose
APPLICANT: Renard, Patricia
APPLICANT: Art, Muriel
TITLE OF INVENTION: METHOD AND KIT FOR THE DETERMINATION OF
FILE REFERENCE: VANM212.001CPI
CURRENT APPLICATION NUMBER: US/10/339,161
CURRENT FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: US 09/816,763
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: EP 00870057.7
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatsSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 19
```

```
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Transcription factor p53
US-10-339-161-14
```

```
Query Match
Best Local Similarity 100.0%; Score 19; DB 15; Length 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGACATGCCCGGGCATGTTC 19
DB 1 GGACATGCCCGGGCATGTTC 19
```

```
RESULT 8
US-10-339-161-14/c
Sequence 14, Application US/10339161
Publication No. US20030162211A1
GENERAL INFORMATION:
APPLICANT: Remacle, Jose
APPLICANT: Art, Muriel
TITLE OF INVENTION: METHOD AND KIT FOR THE DETERMINATION OF
FILE REFERENCE: VANM212.001CPI
CURRENT APPLICATION NUMBER: US/10/339,161
CURRENT FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: US 09/816,763
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: EP 00870057.7
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatsSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 19
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Transcription factor p53
US-10-339-161-14
```

```
Query Match
Best Local Similarity 95.0%; Score 19; DB 15; Length 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 GACATGCCCGGGCATGTCC 20
DB 19 GACATGCCCGGGCATGTCC 1
```

```
RESULT 9
US-10-767-701-2482
Sequence 2482, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 2482
LENGTH: 729
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS41949_1
US-10-767-701-2482
```

Query Match 87.0%; Score 17.4; DB 17; Length 729;
Best Local Similarity 94.7%; Pred. No. 34;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTC 19
DB 247 GGACACGCCCGGGCATGTC 265

RESULT 10
US-10-767-701-2482/c

; Sequence 2482, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 2482
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS41949_1
US-10-767-701-2482

Query Match 87.0%; Score 17.4; DB 17; Length 729;
Best Local Similarity 94.7%; Pred. No. 34;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGGCATGTC 20
DB 265 GACATGCCCGGGCATGTC 247

RESULT 11
US-10-425-114-30682
; Sequence 30682, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30682
; LENGTH: 3325
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73063H08_F11
US-10-425-114-30682

Query Match 87.0%; Score 17.4; DB 16; Length 3325;
Best Local Similarity 94.7%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTC 19
DB 739 GGACACGCCCGGGCATGTC 757

RESULT 12
US-10-425-114-30682/c

; Sequence 30682, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30682
; LENGTH: 3325
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73063H08_F11
US-10-425-114-30682

Query Match 87.0%; Score 17.4; DB 16; Length 3325;
Best Local Similarity 94.7%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGGCATGTC 20
DB 757 GACATGCCCGGGCATGTC 739

RESULT 13
US-10-425-115-4271
; Sequence 4271, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 4271
; LENGTH: 3458
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_103895C.1
US-10-425-115-4271

Query Match 87.0%; Score 17.4; DB 18; Length 3458;
Best Local Similarity 94.7%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTC 19
DB 739 GGACACGCCCGGGCATGTC 757

RESULT 14
US-10-425-115-4271/c
; Sequence 4271, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 4271
LENGTH: 3458
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MFT4577_103895C.1
US-10-425-115-4271

Query Match 87.0%; Score 17.4; DB 18; Length 3458;
Best Local Similarity 94.7%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCGCATGTC 20
Db 757 GACATGCCCGGCGCGTGTCC 739

RESULT 15
US-10-425-114-2765
Sequence 2765, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 2765
LENGTH: 3569
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700223564_FLI
US-10-425-114-2765

Query Match 87.0%; Score 17.4; DB 16; Length 3569;
Best Local Similarity 94.7%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 19
Db 733 GGACATGCCCGGCGCATGTC 751

RESULT 16
US-10-425-114-2765/C
Sequence 2765, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 2765
LENGTH: 3569
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700223564_FLI
US-10-425-114-2765

FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 2765
LENGTH: 3569
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700223564_FLI
US-10-425-114-2765

Query Match 87.0%; Score 17.4; DB 16; Length 3569;
Best Local Similarity 94.7%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCGCATGTC 20
Db 751 GACATGCCCGGCGCGTGTCC 733

RESULT 17
US-10-425-115-4270
Sequence 4270, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 4270
LENGTH: 3609
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MFT4577_103894C.1
US-10-425-115-4270

Query Match 87.0%; Score 17.4; DB 18; Length 3609;
Best Local Similarity 94.7%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 19
Db 755 GGACATGCCCGGCGCATGTC 773

RESULT 18
US-10-425-115-4270/C
Sequence 4270, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 4270
LENGTH: 3609
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700223564_FLI
US-10-425-115-4270

OTHER INFORMATION: Clone ID: MNT4577_103894C.1
US-10-425-115-4270

Query Match 87.0%; Score 17.4; DB 18; Length 3609;
Best Local Similarity 94.7%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGGCATGTCC 20
DB 773 GACATGCCCGGGCATGTCC 755

RESULT 19
US-09-829-922-31

Sequence 31, Application US/09829922
Publication No. US20030171537A1

GENERAL INFORMATION:

APPLICANT: Halazonec's, Thanos

APPLICANT: Hartwig, Wolfgang

TITLE OF INVENTION: Peptides and peptidomimetics with

TITLE OF INVENTION: structural similarity to human p53 that activate p53

TITLE OF INVENTION: function

FILE REFERENCE: 2973.19998

CURRENT APPLICATION NUMBER: US/09/829,922

CURRENT FILING DATE: 2001-04-11

PRIOR APPLICATION NUMBER: 08/894,327

PRIOR FILING DATE: 1997-12-04

PRIOR APPLICATION NUMBER: pctus96/01535

PRIOR FILING DATE: 1996-02-16

PRIOR APPLICATION NUMBER: 08/392,542

PRIOR FILING DATE: 1995-02-16

NUMBER OF SEQ ID NOS: 35

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 31

LENGTH: 30

TYPE: DNA

ORGANISM: Homo sapiens

US-09-829-922-31

Query Match 84.0%; Score 16.8; DB 10; Length 30;

Best Local Similarity 90.0%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20
DB 3 GGACATGTCCGGGCATGTCC 22

RESULT 20
US-09-829-922-31/c

Sequence 31, Application US/09829922
Publication No. US20030171537A1

GENERAL INFORMATION:

APPLICANT: Halazonec's, Thanos

APPLICANT: Hartwig, Wolfgang

TITLE OF INVENTION: Peptides and peptidomimetics with

TITLE OF INVENTION: structural similarity to human p53 that activate p53

TITLE OF INVENTION: function

FILE REFERENCE: 2973.19998

CURRENT APPLICATION NUMBER: US/09/829,922

CURRENT FILING DATE: 2001-04-11

PRIOR APPLICATION NUMBER: 08/894,327

PRIOR FILING DATE: 1997-12-04

PRIOR APPLICATION NUMBER: pctus96/01535

PRIOR FILING DATE: 1996-02-16

PRIOR APPLICATION NUMBER: 08/392,542

PRIOR FILING DATE: 1995-02-16

NUMBER OF SEQ ID NOS: 35

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 31

LENGTH: 30

TYPE: DNA

ORGANISM: Homo sapiens

US-09-829-922-31

Query Match 84.0%; Score 16.8; DB 10; Length 30;
Best Local Similarity 90.0%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20
DB 22 GGACATGCCCGGGCATGTCC 3

RESULT 21
US-10-160-290-31

Sequence 31, Application US/10160290
Publication No. US2003012457A1

GENERAL INFORMATION:

APPLICANT: Halazonec's, Thanos

APPLICANT: Hartwig, Wolfgang

TITLE OF INVENTION: Peptides and peptidomimetics with

TITLE OF INVENTION: structural similarity to Human p53 That Activate

TITLE OF INVENTION: function

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner, Birch, McKie & Beckett

STREET: 1001 G Street, N.W.

CITY: Washington, D.C.

STATE: District of Columbia

COUNTRY: U.S.

ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/160,290

FILING DATE: 04-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/685,027

FILING DATE: 10-Oct-2000

APPLICATION NUMBER: 08/392,542

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Boborske, Laurence H.

REGISTRATION NUMBER: 34,698

REFERENCE/DOCKET NUMBER: 0486,48439

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202 508-9100

TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-10-160-290-31

Query Match 84.0%; Score 16.8; DB 15; Length 30;
Best Local Similarity 90.0%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTCC 20
DB 3 GGACATGTCCGGGCATGTCC 22

RESULT 22

US-10-160-290-31/c
; Sequence 31, Application US/10160290
; Publication No. US20030124557A1
; GENERAL INFORMATION:
; APPLICANT: Halazonetis, Thanos
; TITLE OF INVENTION: Peptides nad Peptidomimetics with
; Structural Similarity to Human p53 That Activate
; p53
; Function
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/160,290
; FILING DATE: 04-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/685,027
; FILING DATE: 10-Oct-2000
; APPLICATION NUMBER: 08/392,542
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 0486,48439
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9100
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-10-160-290-31
Query Match 84.0%; Score 16.8; DB 15; Length 30;
Best Local Similarity 90.0%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Cy 1 GGACATGCCCGGCGCATGTCC 20
Db 22 GGACATGCCCGGCGCATGTCC 3
RESULT 23
US-10-052-482-61
; Sequence 61, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 61
; LENGTH: 21565
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1217)..(1236)
; OTHER INFORMATION: "n" at positions 1217 to 1236 can be any base
; NAME/KEY: misc_feature
; LOCATION: (3121)..(3140)
; OTHER INFORMATION: "n" at positions 3121 to 3140 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5706)..(5725)
; OTHER INFORMATION: "n" at positions 5706 to 5725 can be any base
US-10-052-482-61
Query Match 84.0%; Score 16.8; DB 16; Length 21565;
Best Local Similarity 90.0%; Pred. No. 45;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Cy 1 GGACATGCCCGGCGCATGTCC 20
Db 9378 GGCGTTCGCCGGCGCATGTCC 9378
RESULT 24
US-10-052-482-61/c
; Sequence 61, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 61
; LENGTH: 21565
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1217)..(1236)
; OTHER INFORMATION: "n" at positions 1217 to 1236 can be any base
; NAME/KEY: misc_feature
; LOCATION: (3121)..(3140)
; OTHER INFORMATION: "n" at positions 3121 to 3140 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5706)..(5725)
; OTHER INFORMATION: "n" at positions 5706 to 5725 can be any base
US-10-052-482-61
Query Match 84.0%; Score 16.8; DB 16; Length 21565;
Best Local Similarity 90.0%; Pred. No. 45;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Cy 1 GGACATGCCCGGCGCATGTCC 20
Db 9397 GGCGTTCGCCGGCGCATGTCC 9378

```
RESULT 25
US-10-367-094-1
; Sequence 1, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 46030
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(46030)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-1

Query Match      84.0%; Score 16.8; DB 17; Length 46030;
Best Local Similarity 90.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGACATGCCCGGCGCATGTCC 20
Db      6907 GGGCTGTGCCCGGCGCATGTCC 6926

RESULT 26
US-10-367-094-1/c
; Sequence 1, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 46030
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(46030)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-1

Query Match      84.0%; Score 16.8; DB 17; Length 46030;
Best Local Similarity 90.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGACATGCCCGGCGCATGTCC 20
Db      6926 GGACATGCCCGGCGCATGTCC 6907
```

```
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 187
; LENGTH: 60327
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1943)..(1968)
; OTHER INFORMATION: "n" at positions 1949 to 1968 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (8743)..(8762)
; OTHER INFORMATION: "n" at positions 8743 to 8762 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (14521)..(14618)
; OTHER INFORMATION: "n" at positions 14521 to 14618 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (18936)..(18955)
; OTHER INFORMATION: "n" at positions 18936 to 18955 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (25819)..(26117)
; OTHER INFORMATION: "n" at positions 25819 to 26117 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (38774)..(39110)
; OTHER INFORMATION: "n" at positions 38774 to 39110 can be any base
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (48067)..(48086)
; OTHER INFORMATION: "n" at positions 48067 to 48086 can be any base
US-10-052-482-187

Query Match      84.0%; Score 16.8; DB 16; Length 60327;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGACATGCCCGGCGCATGTCC 20
Db      3740 GGACATGCCCGGCGCATGTCC 3759

RESULT 28
US-10-052-482-187/c
; Sequence 187, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 187
; LENGTH: 60327
; TYPE: DNA
```

```

; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1945)..(1968)
; OTHER INFORMATION: "n" at positions 1949 to 1968 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8743)..(8762)
; OTHER INFORMATION: "n" at positions 8743 to 8762 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14521)..(14618)
; OTHER INFORMATION: "n" at positions 14521 to 14618 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18936)..(18955)
; OTHER INFORMATION: "n" at positions 18936 to 18955 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (25819)..(26117)
; OTHER INFORMATION: "n" at positions 25819 to 26117 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (38774)..(39110)
; OTHER INFORMATION: "n" at positions 38774 to 39110 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (48067)..(48086)
; OTHER INFORMATION: "n" at positions 48067 to 48086 can be any base
US-10-052-482-187
```

```

Query Match      84.0%; Score 16.8; DB 16; Length 60327;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

Qy      1 GGACATGCCCGGGCATGTCC 20
        |||
Db      3759 GGGCCTGCCCGGGCATGTCC 3740
```

```

RESULT 29
US-10-034-650-31
; Sequence 31, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000128
; CURRENT APPLICATION NUMBER: US/10/034,650
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/474,377
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 68233
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-034-650-31
```

```

Query Match      84.0%; Score 16.8; DB 15; Length 68233;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

Qy      1 GGACATGCCCGGGCATGTCC 20
        |||
Db      14433 GGACATGCCCGGGCATGTCC 14452
```

```

RESULT 30
US-10-034-650-31/c
; Sequence 31, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000128
; CURRENT APPLICATION NUMBER: US/10/034,650
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/474,377
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 68233
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-034-650-31
```

```

Query Match      84.0%; Score 16.8; DB 15; Length 68233;
Best Local Similarity 90.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

Qy      1 GGACATGCCCGGGCATGTCC 20
        |||
Db      14452 GGGCCTGCCCGGGCATGTCC 14433
```

```

RESULT 31
US-10-087-192-709
; Sequence 709, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 709
; LENGTH: 77941
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..((77941))
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-709
```

```

Query Match      84.0%; Score 16.8; DB 13; Length 77941;
Best Local Similarity 90.0%; Pred. No. 39;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

Qy      1 GGACATGCCCGGGCATGTCC 20
        |||
Db      65117 GGGCCTGCCCGGGCATGTCC 65136
```

```

RESULT 32
US-10-087-192-709/c
; Sequence 709, Application US/10087192
```



```
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: CANCER
FILE REFERENCE: 52945200122
CURRENT APPLICATION NUMBER: US/10/087,192
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 709
LENGTH: 77941
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(77941)
OTHER INFORMATION: n = A,T,C or G
US-10-087-192-709

Query Match      84.0%; Score 16.8; DB 13; Length 77941;
Best Local Similarity 90.0%; Pred. No. 39;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 GGACATGCCCGGCGCATGTCC 20
Db      65136 GGACATGCCCGGCGCATGTCC 65117

RESULT 33
US-10-322-281-33
Sequence 33, Application US/10322281
Publication No. US20040126762A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
APPLICANT: Marc S. Malandro
TITLE OF INVENTION: Novel Compositions and Methods in Cancer
FILE REFERENCE: 52945200100
CURRENT APPLICATION NUMBER: US/10/322,281
CURRENT FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 866
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 145068
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(145068)
OTHER INFORMATION: n = A,T,C or G
US-10-322-281-33

Query Match      84.0%; Score 16.8; DB 17; Length 145068;
Best Local Similarity 90.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 GGACATGCCCGGCGCATGTCC 20
Db      95566 GGACATGCCCGGCGCATGTCC 95565

RESULT 34
US-10-322-281-33/c
Sequence 33, Application US/10322281
Publication No. US20040126762A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
APPLICANT: Marc S. Malandro
```

```
TITLE OF INVENTION: Novel Compositions and Methods in Cancer
FILE REFERENCE: 52945200100
CURRENT APPLICATION NUMBER: US/10/322,281
CURRENT FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 866
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 145068
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(145068)
OTHER INFORMATION: n = A,T,C or G
US-10-322-281-33

Query Match      84.0%; Score 16.8; DB 17; Length 145068;
Best Local Similarity 90.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 GGACATGCCCGGCGCATGTCC 20
Db      95565 GGACATGCCCGGCGCATGTCC 95566

RESULT 35
US-09-835-976B-77
Sequence 77, Application US/09835976B
Publication No. US20030027983A1
GENERAL INFORMATION:
APPLICANT: Mount, David B.
APPLICANT: Delpire, Eric
APPLICANT: Gamba, Gerardo
APPLICANT: Alfred L. George, Jr.
TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
TITLE OF INVENTION: POLYPEPTIDES AND
TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
CURRENT APPLICATION NUMBER: US/09/835,976B
CURRENT FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 131
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 77
LENGTH: 400
TYPE: DNA
ORGANISM: Homo sapiens
US-09-835-976B-77

Query Match      82.0%; Score 16.4; DB 10; Length 400;
Best Local Similarity 94.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      3 ACATGCCCGGCGCATGTCC 20
Db      351 ACATGCCCGGCGCATGTCC 368

RESULT 36
US-09-835-976B-77/c
Sequence 77, Application US/09835976B
Publication No. US20030027983A1
GENERAL INFORMATION:
APPLICANT: Mount, David B.
APPLICANT: Delpire, Eric
APPLICANT: Gamba, Gerardo
APPLICANT: Alfred L. George, Jr.
TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
TITLE OF INVENTION: POLYPEPTIDES AND
TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
CURRENT APPLICATION NUMBER: US/09/835,976B
CURRENT FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 131
```

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 77
LENGTH: 400
TYPE: DNA
ORGANISM: Homo sapiens
US-09-835-976B-77

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 10; Length 400;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGT 18
DB 368 GGACATGCCCGGCATGT 351

RESULT 37
US-09-835-976B-111
Sequence 111, Application US/09835976B
Publication No. US20030027983A1
GENERAL INFORMATION:
APPLICANT: Mount, David B.
APPLICANT: Delplre, Eric
APPLICANT: Gamba, Gerardo
APPLICANT: Alfred L. George, Jr.
TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
FILE REFERENCE: ATTORNEY DOCKET NO. US20030027983A1 1242-26-2
CURRENT FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 131
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 111
LENGTH: 3422
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY:
LOCATION:
OTHER INFORMATION: n=a, c, g, or t
US-09-835-976B-111

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 10; Length 3422;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACATGCCCGGCATGTCC 20
DB 3051 ACATGCCCGGCATGTCC 3068

RESULT 38
US-09-835-976B-111/c
Sequence 111, Application US/09835976B
Publication No. US20030027983A1
GENERAL INFORMATION:
APPLICANT: Mount, David B.
APPLICANT: Delplre, Eric
APPLICANT: Gamba, Gerardo
APPLICANT: Alfred L. George, Jr.
TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
FILE REFERENCE: ATTORNEY DOCKET NO. US20030027983A1 1242-26-2
CURRENT FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 131
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 111
LENGTH: 3422
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY:
LOCATION:
OTHER INFORMATION: n=a, c, g, or t
US-09-835-976B-111

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 10; Length 3422;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGT 18
DB 3068 GGACATGCCCGGCATGT 3051

RESULT 39
US-10-322-696-40
Sequence 40, Application US/10322696
Publication No. US20040166490A1
GENERAL INFORMATION:
APPLICANT: Malandro, Marc
APPLICANT: Morris, David W.
TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
FILE REFERENCE: 529452001200
CURRENT APPLICATION NUMBER: US/10/322,696
CURRENT FILING DATE: 2003-10-17
NUMBER OF SEQ ID NOS: 186
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 40
LENGTH: 32351
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(32351)
OTHER INFORMATION: n = A,T,C or G
US-10-322-696-40

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 17; Length 32351;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACATGCCCGGCATGTCC 20
DB 322 ACATGCCCGGCATGTCC 339

RESULT 40
US-10-322-696-40/c
Sequence 40, Application US/10322696
Publication No. US20040166490A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Malandro, Marc
TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
FILE REFERENCE: 529452001200
CURRENT APPLICATION NUMBER: US/10/322,696
CURRENT FILING DATE: 2003-10-17
NUMBER OF SEQ ID NOS: 186
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 40
LENGTH: 32351
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(32351)
OTHER INFORMATION: n = A,T,C or G
US-10-322-696-40

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 17; Length 32351;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGT 18
|||
Db 339 GGACATGCCCGGCATGT 322

RESULT 41
US-10-087-192-988
; Sequence 988, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 988
; LENGTH: 203264
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-988

Query Match 82.0%; Score 16.4; DB 13; Length 203264;
Best Local Similarity 94.4%; Pred. No. 55;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACATGCCCGGCATGTCC 20
|||
Db 42121 ACATGCCCGGCATGTCC 42138

RESULT 42
US-10-087-192-988/c
; Sequence 988, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 988
; LENGTH: 203264
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-988

Query Match 82.0%; Score 16.4; DB 13; Length 203264;
Best Local Similarity 94.4%; Pred. No. 55;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGT 18
|||
Db 42138 GGACATGCCCGGCATGT 42121

RESULT 43

US-09-294-093B-1681
; Sequence 1681, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 1681
; LENGTH: 274
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700344792H1
US-09-294-093B-1681

Query Match 79.0%; Score 15.8; DB 9; Length 274;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCATGTTC 19
|||
Db 189 GGACATGCCCGGCATGTTC 207

RESULT 44
US-09-294-093B-1681/c
; Sequence 1681, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 1681
; LENGTH: 274
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700344792H1
US-09-294-093B-1681

Query Match 79.0%; Score 15.8; DB 9; Length 274;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCATGTCC 20
|||
Db 207 GACATGCCCGGCATGTCC 189

RESULT 45
US-09-991-936-580
; Sequence 580, Application US/09991936
; Publication No. US20030073827A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.

```

; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 580
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-580

```

```

Query Match      79.0%; Score 15.8; DB 10; Length 277;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GGACATGCCCGGCATGTC 19
DB 129 GGACATTCCTCGGCATGTC 147

```

```

RESULT 46
US-09-991-936-580/c
; Sequence 580, Application US/09991936
; Publication No. US20030073827A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 580
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-580

```

```

Query Match      79.0%; Score 15.8; DB 10; Length 277;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 GACATGCCCGGCATGTC 20
DB 147 GACATGCCCGGCATGTC 129

```

```

RESULT 47
US-09-294-093B-5575
; Sequence 5575, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL

```

```

; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 5575
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700357426H1
; LOCATION: 69
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-5575

```

```

Query Match      79.0%; Score 15.8; DB 9; Length 294;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GGACATGCCCGGCATGTC 19
DB 78 GGACATTCCTCGGCATGTC 96

```

```

RESULT 48
US-09-294-093B-5575/c
; Sequence 5575, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 5575
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700357426H1
; LOCATION: 69
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-5575

```

```

Query Match      79.0%; Score 15.8; DB 9; Length 294;
Best Local Similarity 89.5%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 GACATGCCCGGCATGTC 20
DB 96 GACATGCCCGGCATGTC 78

```

```

RESULT 49
US-09-783-590-4831
; Sequence 4831, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong

```

```

; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4831
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (171)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (218)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (252)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (288)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (333)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (353)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (357)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (417)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (447)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (466)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (468)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (474)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-4831

Query Match          79.0% Score 15.8; DB 9; Length 482;
Best Local Similarity 89.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGACATGCCCGGCGCATGTC 19
Db      148 GGACATGCCCGGCGCATGCC 166
```

```

; Patent No.: US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4831
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (171)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (218)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (252)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (288)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (333)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (353)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (357)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (417)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (447)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (466)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (468)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (474)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-4831

Query Match          79.0% Score 15.8; DB 9; Length 482;
Best Local Similarity 89.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GGACATGCCCGGCGCATGTC 20
Db      166 GGACATGCCCGGCGCATGCC 148
```

Sun Jan 16 16:27:32 2005

us-09-578-453-2.rnpb

Page 18

Search completed: January 14, 2005, 19:48:25
Job time : 381.421 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 14:35:09 ; Search time 2632.63 Seconds
(without alignments)
276.831 Million cell updates/sec

Title: US-09-578-453-2

Perfect score: 20

Sequence: 1 GGACATGCCCGGCGCATGTC 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	92.0	683	9	CR249440 Forward s
2	18.4	92.0	683	9	CR249440 Forward s
3	17.4	87.0	409	8	BH781927 f2mb001f0
4	17.4	87.0	409	8	BH781927 f2mb001f0
5	17.4	87.0	482	4	BG556896 EM1_40_AO
6	17.4	87.0	482	4	BG556896 EM1_40_AO
7	17.4	87.0	557	4	BI098283 IP1_30_D1
8	17.4	87.0	557	4	BI098283 IP1_30_D1
9	17.4	87.0	757	9	CG270390 CG3D059TV
10	17.4	87.0	757	9	CG270390 CG3D059TV
11	17.4	87.0	843	9	CG290290 CG2B047TH
12	17.4	87.0	843	9	CG290290 CG2B047TH
13	17.4	87.0	870	9	CG236275 OG1CU67TH
14	17.4	87.0	870	9	CG236275 OG1CU67TH
15	17.4	87.0	876	8	BZ364970 PUDS823TD
16	17.4	87.0	876	8	BZ364970 PUDS823TD
17	17.4	87.0	877	8	BZ364970 PUDS823TD
18	17.4	87.0	877	8	BZ364970 PUDS823TD
19	17.4	87.0	877	9	CL637471 CH243-SH5
20	17.4	87.0	877	9	CL637471 CH243-SH5
21	17.4	87.0	930	9	CG236287 OG1CU67TV
22	17.4	87.0	930	9	CG236287 OG1CU67TV
23	16.8	84.0	977	9	CG290302 CG2B047TV
24	16.8	84.0	179	9	CL640722 M003D05 G

25	16.8	84.0	221	8	BZ691000 M003D05 G
26	16.8	84.0	221	8	BZ691000 M003D05 G
27	16.8	84.0	227	1	AV322349 AV322349
28	16.8	84.0	227	1	AV322349 AV322349
29	16.8	84.0	269	8	AZ377623 1M0132L01
30	16.8	84.0	269	8	AZ377623 1M0132L01
31	16.8	84.0	274	8	AZ505789 1M0346R12
32	16.8	84.0	274	8	AZ505789 1M0346R12
33	16.8	84.0	280	5	BP090312 BP090312
34	16.8	84.0	280	5	BP090312 BP090312
35	16.8	84.0	282	2	BB093079 BB093079
36	16.8	84.0	282	2	BB093079 BB093079
37	16.8	84.0	283	2	BB157894 BB157894
38	16.8	84.0	283	2	BB157894 BB157894
39	16.8	84.0	283	2	BB755436 BB755436
40	16.8	84.0	283	2	BB755436 BB755436
41	16.8	84.0	287	5	BY469224 BY469224
42	16.8	84.0	287	5	BY469224 BY469224
43	16.8	84.0	288	7	CK337629 CK337629
44	16.8	84.0	288	7	CK337629 CK337629
45	16.8	84.0	288	7	CK340780 CK340780
46	16.8	84.0	288	7	CK340780 CK340780
47	16.8	84.0	290	2	BB060455 BB060455
48	16.8	84.0	290	2	BB060455 BB060455
49	16.8	84.0	293	1	AA289954 AA289954
50	16.8	84.0	293	1	AA289954 AA289954
51	16.8	84.0	303	6	CD541539 CD541539
52	16.8	84.0	303	6	CD541539 CD541539
53	16.8	84.0	306	6	CD539301 CD539301
54	16.8	84.0	306	6	CD539301 CD539301
55	16.8	84.0	307	2	AM213731 AM213731
56	16.8	84.0	307	2	AM213731 AM213731
57	16.8	84.0	311	2	BB154112 BB154112
58	16.8	84.0	311	2	BB154112 BB154112
59	16.8	84.0	319	2	BB132471 BB132471
60	16.8	84.0	319	2	BB132471 BB132471
61	16.8	84.0	321	1	AA408122 AA408122
62	16.8	84.0	321	1	AA408122 AA408122
63	16.8	84.0	324	7	R61912 R61912
64	16.8	84.0	324	7	R61912 R61912
65	16.8	84.0	327	2	BF452129 BF452129
66	16.8	84.0	327	2	BF452129 BF452129
67	16.8	84.0	333	1	AA445841 AA445841
68	16.8	84.0	333	1	AA445841 AA445841
69	16.8	84.0	333	7	CK337615 CK337615
70	16.8	84.0	333	7	CK337615 CK337615
71	16.8	84.0	338	7	CK337751 CK337751
72	16.8	84.0	338	7	CK337751 CK337751
73	16.8	84.0	346	1	AA617393 AA617393
74	16.8	84.0	346	1	AA617393 AA617393
75	16.8	84.0	347	2	BB670724 BB670724
76	16.8	84.0	347	2	BB670724 BB670724
77	16.8	84.0	351	8	AQ992743 RPCI-23-2
78	16.8	84.0	351	8	AQ992743 RPCI-23-2
79	16.8	84.0	358	1	AV117495 AV117495
80	16.8	84.0	358	1	AV117495 AV117495
81	16.8	84.0	361	2	BB802219 BB802219
82	16.8	84.0	361	2	BB802219 BB802219
83	16.8	84.0	366	2	BB809747 BB809747
84	16.8	84.0	366	2	BB809747 BB809747
85	16.8	84.0	369	1	A1427020 A1427020
86	16.8	84.0	369	1	A1427020 A1427020
87	16.8	84.0	376	8	AZ483669 AZ483669
88	16.8	84.0	376	8	AZ483669 AZ483669
89	16.8	84.0	377	1	A1550171 A1550171
90	16.8	84.0	377	1	A1550171 A1550171
91	16.8	84.0	379	2	AM323086 AM323086
92	16.8	84.0	379	2	AM323086 AM323086
93	16.8	84.0	382	2	AM741148 AM741148
94	16.8	84.0	382	2	AM741148 AM741148
95	16.8	84.0	382	2	BB793888 BB793888
96	16.8	84.0	382	2	BB793888 BB793888
97	16.8	84.0	390	2	BB800634 BB800634

C 98	16.8	84.0	390 2	BB800634	BB800634	BB800634	171	16.8	84.0	459 1	A1159253	A1159253 v288a11.r
C 99	16.8	84.0	390 7	CK341151	CK341151	CK341151	C 172	16.8	84.0	459 1	A1159253	A1159253 v288a11.r
C 100	16.8	84.0	390 7	CK341151	CK341151	CK341151	C 173	16.8	84.0	459 6	C87458	C87458 C87458 Mous
C 101	16.8	84.0	391 7	CP947326	CP947326	CP947326	C 174	16.8	84.0	462 6	C87458	C87458 C87458 Mous
C 102	16.8	84.0	391 7	CP947326	CP947326	CP947326	C 175	16.8	84.0	462 9	CL631989	CL631989 C87458 Mous
C 103	16.8	84.0	391 8	AZ240326	AZ240326	AZ240326	C 176	16.8	84.0	462 9	CL631989	CL631989 C87458 Mous
C 104	16.8	84.0	391 8	AZ240326	AZ240326	AZ240326	C 177	16.8	84.0	465 2	BE369064	BE369064 601221734
C 105	16.8	84.0	393 7	CK340826	CK340826	CK340826	C 178	16.8	84.0	465 2	BE369064	BE369064 601221734
C 106	16.8	84.0	393 7	CK340826	CK340826	CK340826	C 179	16.8	84.0	465 7	CK340771	CK340771 K0108507-
C 107	16.8	84.0	394 1	AA125901	AA125901	AA125901	C 180	16.8	84.0	465 7	CK340771	CK340771 K0108507-
C 108	16.8	84.0	394 1	AA125901	AA125901	AA125901	C 181	16.8	84.0	466 8	AZ646335	AZ646335 IM0512L18
C 109	16.8	84.0	394 7	CK342267	CK342267	CK342267	C 182	16.8	84.0	466 8	AZ646335	AZ646335 IM0512L18
C 110	16.8	84.0	394 7	CK342267	CK342267	CK342267	C 183	16.8	84.0	467 1	A1592934	A1592934 vub7c05.x
C 111	16.8	84.0	397 8	AZ788367	AZ788367	AZ788367	C 184	16.8	84.0	467 1	A1592934	A1592934 vub7c05.x
C 112	16.8	84.0	397 8	AZ788367	AZ788367	AZ788367	C 185	16.8	84.0	467 6	CA667592	CA667592 wleui.pk0
C 113	16.8	84.0	398 6	CD548928	CD548928	CD548928	C 186	16.8	84.0	467 6	CA667592	CA667592 wleui.pk0
C 114	16.8	84.0	398 6	CD548928	CD548928	CD548928	C 187	16.8	84.0	468 8	AZ031229	AZ031229 RPT-23-3
C 115	16.8	84.0	399 1	A1450599	A1450599	A1450599	C 188	16.8	84.0	468 8	AZ031229	AZ031229 RPT-23-3
C 116	16.8	84.0	399 1	A1450599	A1450599	A1450599	C 189	16.8	84.0	470 2	BB700662	BB700662 BB700662
C 117	16.8	84.0	402 2	BB802309	BB802309	BB802309	C 190	16.8	84.0	470 2	BB700662	BB700662 BB700662
C 118	16.8	84.0	402 2	BB802309	BB802309	BB802309	C 191	16.8	84.0	470 2	BB700662	BB700662 BB700662
C 119	16.8	84.0	406 2	AA741532	AA741532	AA741532	C 192	16.8	84.0	470 2	BB700662	BB700662 BB700662
C 120	16.8	84.0	406 2	AA741532	AA741532	AA741532	C 193	16.8	84.0	473 5	BO886865	BO886865 AGENCOURT
C 121	16.8	84.0	406 2	BB742521	BB742521	BB742521	C 194	16.8	84.0	473 5	BO886865	BO886865 AGENCOURT
C 122	16.8	84.0	406 2	BB742521	BB742521	BB742521	C 195	16.8	84.0	477 8	AZ264222	AZ264222 RPT-23-1
C 123	16.8	84.0	407 5	BY546952	BY546952	BY546952	C 196	16.8	84.0	477 8	AZ264222	AZ264222 RPT-23-1
C 124	16.8	84.0	408 2	AA206072	AA206072	AA206072	C 197	16.8	84.0	487 2	BB760838	BB760838 BB760838
C 125	16.8	84.0	408 2	AA206072	AA206072	AA206072	C 198	16.8	84.0	487 2	BB760838	BB760838 BB760838
C 126	16.8	84.0	408 5	BB764927	BB764927	BB764927	C 199	16.8	84.0	489 8	AZ585443	AZ585443 IM0390C09
C 127	16.8	84.0	408 5	BB764927	BB764927	BB764927	C 200	16.8	84.0	489 8	AZ585443	AZ585443 IM0390C09
C 128	16.8	84.0	408 5	BB764927	BB764927	BB764927	C 201	16.8	84.0	492 2	AA470831	AA470831 ha34b04.x
C 129	16.8	84.0	408 7	W77307	W77307	W77307	C 202	16.8	84.0	492 2	AA470831	AA470831 ha34b04.x
C 130	16.8	84.0	408 7	W77307	W77307	W77307	C 203	16.8	84.0	493 8	BO055085	BO055085 RPT-24-3
C 131	16.8	84.0	412 8	BH031966	BH031966	BH031966	C 204	16.8	84.0	493 8	BO055085	BO055085 RPT-24-3
C 132	16.8	84.0	412 8	BH031966	BH031966	BH031966	C 205	16.8	84.0	497 8	AZ989198	AZ989198 2M0272A02
C 133	16.8	84.0	415 2	BB788289	BB788289	BB788289	C 206	16.8	84.0	497 8	AZ989198	AZ989198 2M0272A02
C 134	16.8	84.0	415 2	BB788289	BB788289	BB788289	C 207	16.8	84.0	500 7	CK361147	CK361147 C0126G06-
C 135	16.8	84.0	419 1	AA967383	AA967383	AA967383	C 208	16.8	84.0	500 7	CK361147	CK361147 C0126G06-
C 136	16.8	84.0	419 1	AA967383	AA967383	AA967383	C 209	16.8	84.0	501 8	AZ646304	AZ646304 IM0512F11
C 137	16.8	84.0	422 2	BB671022	BB671022	BB671022	C 210	16.8	84.0	501 8	AZ646304	AZ646304 IM0512F11
C 138	16.8	84.0	422 2	BB671022	BB671022	BB671022	C 211	16.8	84.0	503 8	AZ080197	AZ080197 RPT-23-4
C 139	16.8	84.0	423 9	CNS07EUL	CNS07EUL	CNS07EUL	C 212	16.8	84.0	503 8	AZ080197	AZ080197 RPT-23-4
C 140	16.8	84.0	423 9	CNS07EUL	CNS07EUL	CNS07EUL	C 213	16.8	84.0	512 2	BB755827	BB755827 BB755827
C 141	16.8	84.0	427 2	BB790475	BB790475	BB790475	C 214	16.8	84.0	512 2	BB755827	BB755827 BB755827
C 142	16.8	84.0	427 2	BB790475	BB790475	BB790475	C 215	16.8	84.0	514 2	BB700756	BB700756 BB700756
C 143	16.8	84.0	430 1	AU017612	AU017612	AU017612	C 216	16.8	84.0	514 2	BB700756	BB700756 BB700756
C 144	16.8	84.0	430 1	AU017612	AU017612	AU017612	C 217	16.8	84.0	516 8	AZ587847	AZ587847 IM0395K18
C 145	16.8	84.0	431 2	BB727075	BB727075	BB727075	C 218	16.8	84.0	516 8	AZ587847	AZ587847 IM0395K18
C 146	16.8	84.0	431 2	BB727075	BB727075	BB727075	C 219	16.8	84.0	520 8	AZ031430	AZ031430 RPT-23-3
C 147	16.8	84.0	431 2	BB727075	BB727075	BB727075	C 220	16.8	84.0	520 8	AZ031430	AZ031430 RPT-23-3
C 148	16.8	84.0	431 2	BB727075	BB727075	BB727075	C 221	16.8	84.0	521 2	AM060682	AM060682 RPT-23-2
C 149	16.8	84.0	432 2	BB670907	BB670907	BB670907	C 222	16.8	84.0	521 2	AM060682	AM060682 RPT-23-2
C 150	16.8	84.0	432 2	BB670907	BB670907	BB670907	C 223	16.8	84.0	521 2	AM060682	AM060682 RPT-23-2
C 151	16.8	84.0	435 2	AA763751	AA763751	AA763751	C 224	16.8	84.0	521 2	BB755859	BB755859 BB755859
C 152	16.8	84.0	435 2	AA763751	AA763751	AA763751	C 225	16.8	84.0	522 2	BB755859	BB755859 BB755859
C 153	16.8	84.0	435 2	BB728884	BB728884	BB728884	C 226	16.8	84.0	522 2	BB755859	BB755859 BB755859
C 154	16.8	84.0	435 4	BB728884	BB728884	BB728884	C 227	16.8	84.0	525 4	BB784453	BB784453 602110842
C 155	16.8	84.0	435 4	BB728884	BB728884	BB728884	C 228	16.8	84.0	525 4	BB784453	BB784453 602110842
C 156	16.8	84.0	435 4	BB728884	BB728884	BB728884	C 229	16.8	84.0	525 5	BB784453	BB784453 602110842
C 157	16.8	84.0	435 4	BB728884	BB728884	BB728884	C 230	16.8	84.0	525 5	BB784453	BB784453 602110842
C 158	16.8	84.0	436 2	BB728884	BB728884	BB728884	C 231	16.8	84.0	525 7	BB784453	BB784453 602110842
C 159	16.8	84.0	436 7	CK340892	CK340892	CK340892	C 232	16.8	84.0	525 7	CK340892	CK340892 K0137C05-
C 160	16.8	84.0	436 7	CK340892	CK340892	CK340892	C 233	16.8	84.0	526 8	AZ120597	AZ120597 RPT-23-2
C 161	16.8	84.0	438 6	CD802736	CD802736	CD802736	C 234	16.8	84.0	526 8	AZ120597	AZ120597 RPT-23-2
C 162	16.8	84.0	438 6	CD802736	CD802736	CD802736	C 235	16.8	84.0	527 8	AZ120597	AZ120597 RPT-23-2
C 163	16.8	84.0	440 8	AZ379582	AZ379582	AZ379582	C 236	16.8	84.0	527 8	AZ120597	AZ120597 RPT-23-2
C 164	16.8	84.0	440 8	AZ379582	AZ379582	AZ379582	C 237	16.8	84.0	528 2	BB752621	BB752621 BB752621
C 165	16.8	84.0	441 2	AA762330	AA762330	AA762330	C 238	16.8	84.0	528 2	BB752621	BB752621 BB752621
C 166	16.8	84.0	441 2	AA762330	AA762330	AA762330	C 239	16.8	84.0	533 5	BX524773	BX524773 BX524773
C 167	16.8	84.0	443 2	BB759594	BB759594	BB759594	C 240	16.8	84.0	533 5	BX524773	BX524773 BX524773
C 168	16.8	84.0	443 2	BB759594	BB759594	BB759594	C 241	16.8	84.0	533 8	AZ123002	AZ123002 RPT-23-4
C 169	16.8	84.0	456 1	AA267947	AA267947	AA267947	C 242	16.8	84.0	533 8	AZ123002	AZ123002 RPT-23-4
C 170	16.8	84.0	456 1	AA267947	AA267947	AA267947	C 243	16.8	84.0	536 8	BH031963	BH031963 RPT-24-3

C 244	16.8	84.0	536	8	BH031963	BH031963	RPCI-24-3	317	16.8	84.0	626	9	CR082848	CR082848	Reverse B
C 245	16.8	84.0	537	1	AA389154	AA389154	VB24E05..r	318	16.8	84.0	626	9	CR082848	CR082848	Reverse B
C 246	16.8	84.0	537	1	AA389154	AA389154	VB24E05..r	319	16.8	84.0	629	9	AZ113897	AZ113897	RPCI-23-1
C 247	16.8	84.0	537	9	CR264982	CR264982	Reverse B	320	16.8	84.0	629	9	AZ113897	AZ113897	RPCI-23-1
C 248	16.8	84.0	537	9	CR264982	CR264982	Reverse B	321	16.8	84.0	630	7	CP906196	CP906196	A044AB03-
C 249	16.8	84.0	539	8	AZ264310	AZ264310	RPCI-23-1	322	16.8	84.0	630	7	CP906196	CP906196	A044AB03-
C 250	16.8	84.0	539	8	AZ264310	AZ264310	RPCI-23-1	323	16.8	84.0	630	7	BH065057	BH065057	RPCI-24-3
C 251	16.8	84.0	541	8	AZ651848	AZ651848	IM0522H18	324	16.8	84.0	630	8	BH065057	BH065057	RPCI-24-3
C 252	16.8	84.0	541	8	AZ651848	AZ651848	IM0522H18	325	16.8	84.0	632	2	BE281764	BE281764	601099715
C 253	16.8	84.0	543	7	CK340918	CK340918	K013A06-	326	16.8	84.0	632	2	BE281764	BE281764	601099715
C 254	16.8	84.0	543	7	CK340918	CK340918	K013A06-	327	16.8	84.0	632	8	AZ792696	AZ792696	2M0045P15
C 255	16.8	84.0	543	7	CK341119	CK341119	K014BD03-	328	16.8	84.0	632	8	AZ792696	AZ792696	2M0045P15
C 256	16.8	84.0	543	7	CK341119	CK341119	K014BD03-	329	16.8	84.0	634	8	AZ494009	AZ494009	1M0329C06
C 257	16.8	84.0	545	8	AZ889925	AZ889925	RPCI-24-1	330	16.8	84.0	634	8	AZ494009	AZ494009	1M0329C06
C 258	16.8	84.0	548	9	CR141987	CR141987	Reverse B	331	16.8	84.0	640	8	AZ800701	AZ800701	2M0058D19
C 259	16.8	84.0	548	9	CR141987	CR141987	Reverse B	332	16.8	84.0	640	8	AZ800701	AZ800701	2M0058D19
C 260	16.8	84.0	555	2	BE913551	BE913551	601665888	333	16.8	84.0	641	8	AZ505060	AZ505060	1M0345M02
C 261	16.8	84.0	555	2	BE913551	BE913551	601665888	334	16.8	84.0	641	8	AZ505060	AZ505060	1M0345M02
C 262	16.8	84.0	555	2	BE913551	BE913551	601665888	335	16.8	84.0	641	8	AZ505060	AZ505060	1M0345M02
C 263	16.8	84.0	557	4	BG147317	BG147317	mab39h06..	336	16.8	84.0	641	8	AZ505060	AZ505060	1M0345M02
C 264	16.8	84.0	557	4	BG147317	BG147317	mab39h06..	337	16.8	84.0	644	2	BF148642	BF148642	uy32f09..x
C 265	16.8	84.0	559	1	AA110628	AA110628	mps2801..r	338	16.8	84.0	644	2	BF148642	BF148642	uy32f09..x
C 266	16.8	84.0	559	1	AA110628	AA110628	mps2801..r	339	16.8	84.0	649	2	CR009325	CR009325	Reverse B
C 267	16.8	84.0	559	8	AZ367994	AZ367994	IM0117C21	340	16.8	84.0	649	9	CR009325	CR009325	Reverse B
C 268	16.8	84.0	559	8	AZ367994	AZ367994	IM0117C21	341	16.8	84.0	650	8	AQ985141	AQ985141	RPCI-23-3
C 269	16.8	84.0	560	4	BG067777	BG067777	H3058A11-	342	16.8	84.0	650	8	AQ985141	AQ985141	RPCI-23-3
C 270	16.8	84.0	560	4	BG067777	BG067777	H3058A11-	343	16.8	84.0	650	8	AZ241609	AZ241609	RPCI-23-8
C 271	16.8	84.0	568	8	AZ336285	AZ336285	1M0066019	344	16.8	84.0	650	8	AZ241609	AZ241609	RPCI-23-8
C 272	16.8	84.0	568	8	AZ336285	AZ336285	1M0066019	345	16.8	84.0	650	8	AZ241609	AZ241609	RPCI-23-8
C 273	16.8	84.0	571	6	CA877937	CA877937	K0961F08-	346	16.8	84.0	650	8	AZ241609	AZ241609	RPCI-23-8
C 274	16.8	84.0	571	6	CA877937	CA877937	K0961F08-	347	16.8	84.0	651	8	AZ609056	AZ609056	1M0433C12
C 275	16.8	84.0	575	6	CD876173	CD876173	AZ03..1080	348	16.8	84.0	651	8	AZ609056	AZ609056	1M0433C12
C 276	16.8	84.0	575	6	CD876173	CD876173	AZ03..1080	349	16.8	84.0	655	8	AZ764190	AZ764190	1M0560E05
C 277	16.8	84.0	576	8	AZ790675	AZ790675	2M0039118	350	16.8	84.0	655	8	AZ764190	AZ764190	1M0560E05
C 278	16.8	84.0	576	8	AZ790675	AZ790675	2M0039118	351	16.8	84.0	657	8	AZ791482	AZ791482	2M0041J23
C 279	16.8	84.0	578	8	AZ587544	AZ587544	1M0395K18	352	16.8	84.0	657	8	AZ791482	AZ791482	2M0041J23
C 280	16.8	84.0	578	8	AZ587544	AZ587544	1M0395K18	353	16.8	84.0	657	8	AZ791482	AZ791482	2M0041J23
C 281	16.8	84.0	580	8	AZ473435	AZ473435	1M0289N18	354	16.8	84.0	657	8	AZ505051	AZ505051	2M0214J11
C 282	16.8	84.0	580	8	AZ473435	AZ473435	1M0289N18	355	16.8	84.0	658	8	AZ505051	AZ505051	2M0214J11
C 283	16.8	84.0	581	9	CR262174	CR262174	Forward B	356	16.8	84.0	658	8	AZ505051	AZ505051	2M0214J11
C 284	16.8	84.0	581	9	CR262174	CR262174	Forward B	357	16.8	84.0	660	8	AZ505051	AZ505051	2M0214J11
C 285	16.8	84.0	584	8	AZ327480	AZ327480	IM0050L17	358	16.8	84.0	660	8	AZ505051	AZ505051	2M0214J11
C 286	16.8	84.0	584	8	AZ327480	AZ327480	IM0050L17	359	16.8	84.0	661	1	AV303415	AV303415	AV303415
C 287	16.8	84.0	587	8	AZ355676	AZ355676	1M0059N24	360	16.8	84.0	661	1	AV303415	AV303415	AV303415
C 288	16.8	84.0	587	8	AZ355676	AZ355676	1M0059N24	361	16.8	84.0	661	8	AZ450945	AZ450945	1M0249022
C 289	16.8	84.0	588	8	AQ934840	AQ934840	RPCI-23-2	362	16.8	84.0	661	8	AZ450945	AZ450945	1M0249022
C 290	16.8	84.0	590	6	CA879114	CA879114	K0971H07-	363	16.8	84.0	663	9	CR276027	CR276027	Reverse B
C 291	16.8	84.0	590	6	CA879114	CA879114	K0971H07-	364	16.8	84.0	663	9	CR276027	CR276027	Reverse B
C 292	16.8	84.0	590	6	CA879114	CA879114	K0971H07-	365	16.8	84.0	665	2	BE287465	BE287465	601097054
C 293	16.8	84.0	590	8	AZ326978	AZ326978	1M0050A08	366	16.8	84.0	665	2	BE287465	BE287465	601097054
C 294	16.8	84.0	590	8	AZ326978	AZ326978	1M0050A08	367	16.8	84.0	672	8	AZ581892	AZ581892	1M0374M06
C 295	16.8	84.0	593	8	AZ255622	AZ255622	RPCI-23-1	368	16.8	84.0	672	8	AZ581892	AZ581892	1M0374M06
C 296	16.8	84.0	593	8	AZ255622	AZ255622	RPCI-23-1	369	16.8	84.0	684	8	AZ581892	AZ581892	1M0374M06
C 297	16.8	84.0	594	2	BF143384	BF143384	601790024	370	16.8	84.0	684	8	AZ581892	AZ581892	1M0374M06
C 298	16.8	84.0	594	2	BF143384	BF143384	601790024	371	16.8	84.0	687	6	BY748643	BY748643	2M0244P11
C 299	16.8	84.0	595	8	AZ903084	AZ903084	RPCI-24-1	372	16.8	84.0	687	6	BY748643	BY748643	2M0244P11
C 300	16.8	84.0	595	8	AZ903084	AZ903084	RPCI-24-1	373	16.8	84.0	691	9	CR134687	CR134687	Forward B
C 301	16.8	84.0	600	5	BU919589	BU919589	6032-21 M	374	16.8	84.0	691	9	CR134687	CR134687	Forward B
C 302	16.8	84.0	600	5	BU919589	BU919589	6032-21 M	375	16.8	84.0	693	8	AZ503640	AZ503640	1M0343C05
C 303	16.8	84.0	600	8	AZ609157	AZ609157	1M0433H18	376	16.8	84.0	693	8	AZ503640	AZ503640	1M0343C05
C 304	16.8	84.0	605	5	BQ268631	BQ268631	1156C08..x	377	16.8	84.0	693	9	BX962618	BX962618	Forward B
C 305	16.8	84.0	605	5	BQ268631	BQ268631	1156C08..x	378	16.8	84.0	693	9	BX962618	BX962618	Forward B
C 306	16.8	84.0	609	8	AZ972769	AZ972769	2M0246E16	379	16.8	84.0	694	8	AZ989689	AZ989689	2M0273E08
C 307	16.8	84.0	609	8	AZ972769	AZ972769	2M0246E16	380	16.8	84.0	694	8	AZ989689	AZ989689	2M0273E08
C 308	16.8	84.0	619	8	AZ871987	AZ871987	2M0185P10	381	16.8	84.0	697	8	AZ271298	AZ271298	RPCI-23-1
C 309	16.8	84.0	619	8	AZ871987	AZ871987	2M0185P10	382	16.8	84.0	697	8	AZ271298	AZ271298	RPCI-23-1
C 310	16.8	84.0	619	8	AZ871987	AZ871987	2M0185P10	383	16.8	84.0	699	8	AZ707813	AZ707813	RPCI-23-2
C 311	16.8	84.0	622	8	AZ947859	AZ947859	2M0210P12	384	16.8	84.0	699	8	AZ707813	AZ707813	RPCI-23-2
C 312	16.8	84.0	622	8	AZ947859	AZ947859	2M0210P12	385	16.8	84.0	699	8	AZ982053	AZ982053	2M0262011
C 313	16.8	84.0	625	5	BP765720	BP765720	602892804	386	16.8	84.0	699	8	AZ982053	AZ982053	2M0262011
C 314	16.8	84.0	625	5	BP765720	BP765720	602892804	387	16.8	84.0	707	4	BI105779	BI105779	602892384
C 315	16.8	84.0	626	2	BF168283	BF168283	601773960	388	16.8	84.0	707	4	BI105779	BI105779	602892384
C 316	16.8	84.0	626	2	BF168283	BF168283	601773960	389	16.8	84.0	708	8	AZ379355	AZ379355	1M0334P05

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20
 |||||
 Db 172 GGACATGCCCGGCGCGCATGTCC 191

RESULT 2
 CR249440/c 683 bp DNA linear GSS 06-JUL-2004
 LOCUS CR249440
 DEFINITION Forward strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN63p13, genomic survey sequence.

ACCESSION CR249440
 VERSION CR249440.1 GI:50028293
 KEYWORDS GSS; genome survey sequence; MICEP.
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 683)
 Adams.D.J., Biggs.P.J., Cox.A.V., Davies.R.M., van der Weyden.L., Jonkers.J., Smith.J., Plumb.R.W., Taylor.R.G., Nishijima.I., Yu.Y., Rogers.J. and Bradley.A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICEP

FEATURES
 source
 1..683
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHPN63p13"
 /clone_lib="MHPN"

ORIGIN
 Query Match 92.0%; Score 18.4; DB 9; Length 683;
 Best Local Similarity 95.0%; Pred. No. 2.8e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTCC 20
 |||||
 Db 191 GGACCTGCCCGGCGCATGTCC 172

RESULT 3
 BH781927 409 bp DNA linear GSS 28-MAR-2002
 LOCUS BH781927
 DEFINITION fzm001f001g06k0 fzm filtered library Zea mays genomic clone
 fzm001f001g06 5', genomic survey sequence.

ACCESSION BH781927
 VERSION BH781927.1 GI:19785173
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 409)
 Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
 TITLE Genethresher methylation filtered genomic sequences from maize
 JOURNAL Unpublished (2002)
 COMMENT Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 Plate: fzm001f001 row: g column: 06
 Seq primer: SK reverse
 Classes: Shotgun
 High quality sequence stop: 409.
 Location/Qualifiers
 1..409

FEATURES
 source

/organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="Mo17"
 /db_xref="taxon:4577"
 /clone="fzm001f001g06"
 /clone_lib="fzm filtered library"
 /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells."

ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 409;
 Best Local Similarity 94.7%; Pred. No. 8.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GACATGCCCGGCGCATGTCC 20
 |||||
 Db 45 GACATGCCCGGCGCGCATGTCC 63

RESULT 4
 BH781927/c 409 bp DNA linear GSS 28-MAR-2002
 LOCUS BH781927
 DEFINITION fzm001f001g06k0 fzm filtered library Zea mays genomic clone
 fzm001f001g06 5', genomic survey sequence.

ACCESSION BH781927
 VERSION BH781927.1 GI:19785173
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 409)
 Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
 TITLE Genethresher methylation filtered genomic sequences from maize
 JOURNAL Unpublished (2002)
 COMMENT Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 Plate: fzm001f001 row: g column: 06
 Seq primer: SK reverse
 Classes: Shotgun
 High quality sequence stop: 409.
 Location/Qualifiers
 1..409
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="Mo17"
 /db_xref="taxon:4577"
 /clone="fzm001f001g06"
 /clone_lib="fzm filtered library"
 /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells."

FEATURES

source

ORIGIN
 Query Match 87.0%; Score 17.4; DB 8; Length 409;
 Best Local Similarity 94.7%; Pred. No. 8.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTTC 19
 |||||
 Db 63 GGACACGCCCGGCGCATGTTC 45

RESULT 5
BG556896 482 bp mRNA linear EST 10-APR-2001
LOCUS EMI_40_A07.b1_A002 Embryo 1 (EMI) Sorghum bicolor cDNA, mRNA
DEFINITION BG556896
ACCESSION BG556896
VERSION BG556896.1 GI:13585894
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 482)
Reid,S.P., Cordonnier-Pratt,M.-M., Gingie,A. and Pratt,L.H.
An EST database from Sorghum: developing embryos
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 480
POLYA=No.

FEATURES
source 1..482
location/Qualifiers
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Embryo 1 (EMI)"
/note="Organ: Embryos germinated for 24 hr; Vector: pBluescript II from lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

ORIGIN
Query Match 87.0%; Score 17.4; DB 4; Length 482;
Best Local Similarity 94.7%; Pred. No. 8.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCATGTC 19
|||||
247 GGACACGCCCGGCATGTC 265

RESULT 6
BG556896 482 bp mRNA linear EST 10-APR-2001
LOCUS EMI_40_A07.b1_A002 Embryo 1 (EMI) Sorghum bicolor cDNA, mRNA
DEFINITION BG556896
ACCESSION BG556896
VERSION BG556896.1 GI:13585894
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 482)
Reid,S.P., Cordonnier-Pratt,M.-M., Gingie,A. and Pratt,L.H.
An EST database from Sorghum: developing embryos
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

RESULT 7
BI098283 557 bp mRNA linear EST 26-JUN-2001
LOCUS IP1_30_D12.b1_A002 Immature panicle 1 (IP1) Sorghum bicolor cDNA, mRNA sequence.
DEFINITION BI098283
ACCESSION BI098283
VERSION BI098283.1 GI:14569865
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 557)
Klein,R.R., Cordonnier-Pratt,M.-M., Gingie,A., Sudman,M. and Pratt,L.H.
An EST database from Sorghum: developing preanthesis panicles
Unpublished (2001)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for high quality sequence is 20. Three-prime sequences, which are obtained with PolyTmix or T7 sequencing primer, are presented as the reverse complement.
Seq primer: JEN REV
High quality sequence stop: 551
POLYA=No.

FEATURES
source 1..557
location/Qualifiers
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Immature panicle 1 (IP1)"
/note="Organ: Developing preanthesis panicles; Vector: pBluescript II SK(-) from lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the

ORIGIN
Query Match 87.0%; Score 17.4; DB 4; Length 482;
Best Local Similarity 94.7%; Pred. No. 8.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 GACATGCCCGGCATGCC 20
|||||
265 GACATGCCCGGCATGCC 247

cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."

Query Match 87.0%; Score 17.4; DB 4; Length 557;
Best Local Similarity 94.7%; Pred. No. 8.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 19
|||||
Db 75 GGACATGCCCGGCGCATGTC 93

RESULT 8
BI098283/c
LOCUS
DEFINITION
BIO98283 557 bp mRNA linear EST 26-JUN-2001
Ipi_30_D12_b1_A002 Immature panicle 1 (IPI) Sorghum bicolor cDNA,
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BIO98283.1 GI:14569865
EST.
Sorghum bicolor (sorghum)
Sorghum bicolor (sorghum);
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 557)
Klein, R.R., Cordonnier-Pratt, M.-M., Gingle, A., Sudman, M. and
Pratt, L.H.
An EST database from Sorghum: developing preanthesis panicles
Unpublished (2001)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for high quality sequence is
20. Three-prime sequences, which are obtained with PolyMix or T7
sequencing primer, are presented as the reverse complement.
Seq primer: JEN REV
High quality sequence stop: 551
POLYA=No.

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

FEATURES
source
1. .557
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clone_lib="Immature panicle 1 (IPI)"
/note="Organ: Developing preanthesis panicles; Vector:
pBluescript II SK(-) from lambda ZAP II; Site 1: XhoI;
Site 2: EcoRI; The library was made from poly-A RNA in the
cloning vector lambda ZAP II. Clones to be sequenced were
prepared by mass excision."

ORIGIN
Query Match 87.0%; Score 17.4; DB 4; Length 557;
Best Local Similarity 94.7%; Pred. No. 8.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCGCATGTC 20
|||||
Db 93 GACATGCCCGGCGCATGTC 75

RESULT 9
CG270390
LOCUS
DEFINITION
CG270390 757 bp DNA linear GSS 25-AUG-2003
OG3DQ59TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0783J21,
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CG270390
CG270390.1 GI:34182531
GSS.
Zea mays

REFERENCE
AUTHORS
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 757)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Reenick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunez, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG3DQ59TH
Contact: Cathy Whitelaw
TIGR

TITLE
JOURNAL
COMMENT

FEATURES
source
1. .757
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 757;
Best Local Similarity 94.7%; Pred. No. 8.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGCGCATGTC 20
|||||
Db 243 GACATGCCCGGCGCATGTC 261

RESULT 10
CG270390/c
LOCUS
DEFINITION
OG3DQ59TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0783J21,
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CG270390.1 GI:34182531
GSS.
Zea mays

REFERENCE
AUTHORS
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 757)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Reenick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunez, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG3DQ59TH
Contact: Cathy Whitelaw
TIGR

TITLE
JOURNAL
COMMENT

FEATURES
source
1. .757
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

ACCESSION      genomic survey sequence.
CG290290
VERSION        CG290290.1  GI:34204504
KEYWORDS
SOURCE
ORGANISM       Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
                1 (bases 1 to 843)
REFERENCE
AUTHORS        WhiteLAW,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
                Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
                Citek,R.W., Numburg,A., Robbins,D. and Lakey,N.
                Consortium for Maize Genomics
TITLE          Unpublished (2002)
JOURNAL
COMMENT        Other_GSSs: CG2BC47TV
                Contact: Cathy WhiteLAW
                TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteLAW@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
location/Qualifiers
    1..843
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone_id="ZMMBMA0751G21"
    /clone_1db="ZM 0.7.1.5 KB"
    /note="Vector: pBSCSK-; Site 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

ORIGIN
Query Match      87.0%; Score 17.4; DB 9; Length 843;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy      1 GGACATGCGCGGCATGTC 19
        |||||
Db      542 GGACATCGCCCGGCATGTC 524

RESULT 13
CG236275      870 bp DNA linear GSS 22-AUG-2003
LOCUS        CG236275
DEFINITION   OGIUC67TH_ZM_0.7.1.5_KB Zea mays genomic clone ZMMBMA073K13,
            genomic survey sequence.
ACCESSION    CG236275
VERSION      CG236275.1  GI:34136161
KEYWORDS     GSS.
ORGANISM     Zea mays
                Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
                1 (bases 1 to 870)
REFERENCE
AUTHORS      WhiteLAW,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
                Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
                Citek,R.W., Numburg,A., Robbins,D. and Lakey,N.
                Consortium for Maize Genomics
TITLE        Unpublished (2002)
JOURNAL
COMMENT      Other_GSSs: OGIUC67TV
                Contact: Cathy WhiteLAW
                TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteLAW@tigr.org
Seq primer: TR
Class: sheared ends.

```

FEATURES
source
Location/Qualifiers
1..870
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0733K13"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 87.0%; Score 17.4; DB 9; Length 870;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTC 19
774 GGACACGCCCGGGCATGTC 792

RESULT 14
LOCUS CG236275/c 870 bp DNA linear GSS 22-AUG-2003
DEFINITION OG1CU677H ZM 0.7-1.5_KB Zea mays genomic clone ZMMBMA0733K13,
genomic survey sequence.
ACCESSION CG236275
VERSION CG236275.1 GI:34136161
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 870)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Reinick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Cleck,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OG1CU677H
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: 7F
Class: sheared ends.
Location/Qualifiers
1..870
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0733K13"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 87.0%; Score 17.4; DB 9; Length 870;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGGCATGTC 20
792 GACATGCCCGGGCATGTC 774

RESULT 15
LOCUS B2964970 876 bp DNA linear GSS 25-MAR-2003

DEFINITION PUDGS23TD ZM 0.6-1.0_KB Zea mays genomic clone ZMMBTA183D21,
genomic survey sequence.
ACCESSION B2964970
VERSION B2964970.1 GI:29180280
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 876)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Reinick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennerzen,J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: 7F
Class: sheared ends.
Location/Qualifiers

FEATURES
source
Location/Qualifiers
1..876
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA183D21"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN
Query Match 87.0%; Score 17.4; DB 8; Length 876;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGGCATGTC 19
789 GGACACGCCCGGGCATGTC 807

RESULT 16
LOCUS B2964970/c 876 bp DNA linear GSS 25-MAR-2003
DEFINITION PUDGS23TD ZM 0.6-1.0_KB Zea mays genomic clone ZMMBTA183D21,
genomic survey sequence.
ACCESSION B2964970
VERSION B2964970.1 GI:29180280
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 876)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Reinick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennerzen,J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: 7F
Class: sheared ends.
Location/Qualifiers

FEATURES
source
Location/Qualifiers
1..876
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA183D21"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN
Query Match 87.0%; Score 17.4; DB 8; Length 876;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCGGGCATGTC 20
792 GACATGCCCGGGCATGTC 774

RESULT 15
LOCUS B2964970 876 bp DNA linear GSS 25-MAR-2003

```

source
1. 876
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM8183D21"
/clone_lib="ZM_0.6.1.0_KB"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN
Query Match      87.0%; Score 17.4; DB 8; Length 876;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GACATGCCCGGCATGTC 20
    |||||
Db 807 GACATGCCCGGCATGTC 789

RESULT 17
CL637471      877 bp   DNA      linear   GSS 28-JUN-2004
LOCUS         CH243-SH5.SP6 CHORI-243 Ovis aries genomic clone CH243-SH5, genomic
DEFINITION    survey sequence.
ACCESSION     CL637471
VERSION       CL637471.1 GI:49370767
KEYWORDS      GSS.
SOURCE        Ovis aries (sheep)
ORGANISM      Ovis aries
              Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
              Caprinae; Ovis.
              1 (bases 1 to 877)
              Zhao S., Shetty J., de Jong P., McEwan J.C. and Oddy H.
REFERENCE     The Institute for Genomic Research
AUTHORS       Department of Eukaryotic Genomics
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhaoc@igr.org
              Seq primer: SP6
              Class: BAC ends.

FEATURES
source
Location/Qualifiers
1. 877
/organism="Ovis aries"
/mol_type="genomic DNA"
/strain="Texel breed"
/db_xref="taxon:9940"
/clone="CH243-SH5"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-243"
/notes="Vector: pPARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
The CHORI-243 sheep (M) (Ovis aries) BAC library produced
by Pieter de Jong's lab at CHORI
http://bacpac.chori.org/library.php?id=162"

ORIGIN
Query Match      87.0%; Score 17.4; DB 9; Length 877;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGACATGCCCGGCATGTC 19
    |||||
Db 808 GGCATGCCCGGCATGTC 826

RESULT 18
CL637471/c

```

```

LOCUS         CL637471      877 bp   DNA      linear   GSS 28-JUN-2004
DEFINITION    CH243-SH5.SP6 CHORI-243 Ovis aries genomic clone CH243-SH5, genomic
survey sequence.
ACCESSION     CL637471
VERSION       CL637471.1 GI:49370767
KEYWORDS      GSS.
SOURCE        Ovis aries (sheep)
ORGANISM      Ovis aries
              Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
              Caprinae; Ovis.
              1 (bases 1 to 877)
              Zhao S., Shetty J., de Jong P., McEwan J.C. and Oddy H.
REFERENCE     The Institute for Genomic Research
AUTHORS       Department of Eukaryotic Genomics
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhaoc@igr.org
              Seq primer: SP6
              Class: BAC ends.

FEATURES
source
Location/Qualifiers
1. 877
/organism="Ovis aries"
/mol_type="genomic DNA"
/strain="Texel breed"
/db_xref="taxon:9940"
/clone="CH243-SH5"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-243"
/notes="Vector: pPARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
The CHORI-243 sheep (M) (Ovis aries) BAC library produced
by Pieter de Jong's lab at CHORI
http://bacpac.chori.org/library.php?id=162"

ORIGIN
Query Match      87.0%; Score 17.4; DB 9; Length 877;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GACATGCCCGGCATGTC 20
    |||||
Db 826 GACATGCCCGGCATGCC 808

RESULT 19
CG236287      930 bp   DNA      linear   GSS 22-AUG-2003
LOCUS         OG1CU67TV_ZM_0.7_1.5_KB Zea mays genomic clone ZM8183K13,
DEFINITION    genomic survey sequence.
ACCESSION     CG236287
VERSION       CG236287.1 GI:34136173
KEYWORDS      GSS.
SOURCE        Zea mays
ORGANISM      Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoidae; Andropogoneae; Zea.
              1 (bases 1 to 930)
              Whiteclaw C.A., Quackenbush J., Van Aken S., Uterback T.,
              Reznick A., Fraser C.M., Buddiman M.A., Bedell J.A., Rohlfing T.,
              Citek R.W., Numburg A., Robbins D. and Lakey N.
              Consortium for Maize Genomics
              Unpublished (2002)
              Other GSSs: OG1CU67TV
              Contact: Cathy Whiteclaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843

```


Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..930

FEATURES
source
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0733K13"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 87.0%; Score 17.4; DB 9; Length 930;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACATGCCCCGGGCGATGTC 20
|||
Db 355 GACATGCCCCGGGCGATGTC 373

RESULT 20
CG236287/c 930 bp DNA linear GSS 22-AUG-2003
LOCUS CG236287
DEFINITION OG1CU67TV ZM 0.7.1.5_KB Zea mays genomic clone ZMMBMA0733K13,
genomic survey sequence.
ACCESSION CG236287
VERSION CG236287.1 GI:34136173
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 930)

REFERENCE
AUTHORS White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Frazer, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG1CU67TV
Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..930

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0733K13"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 930;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATGCCCCGGGCGATGTC 19
|||
Db 373 GACATGCCCCGGGCGATGTC 355

RESULT 21
CG290302 977 bp DNA linear GSS 25-AUG-2003
LOCUS CG290302
DEFINITION OG2BC47TV ZM 0.7.1.5_KB Zea mays genomic clone ZMMBMA0751G21,
genomic survey sequence.

ACCESSION CG290302
VERSION CG290302.1 GI:34204516
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 977)

REFERENCE
AUTHORS White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Frazer, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG2BC47TV
Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..977

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0751G21"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 977;
Best Local Similarity 94.7%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATGCCCCGGGCGATGTC 19
|||
Db 813 GACATGCCCCGGGCGATGTC 831

RESULT 22
CG290302 977 bp DNA linear GSS 25-AUG-2003
LOCUS CG290302
DEFINITION OG2BC47TV ZM 0.7.1.5_KB Zea mays genomic clone ZMMBMA0751G21,
genomic survey sequence.

ACCESSION CG290302
VERSION CG290302.1 GI:34204516
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 977)

REFERENCE
AUTHORS White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Frazer, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG2BC47TV
Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA

FEATURES
 source
 Location/Qualifiers
 1. 977
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0751G21"
 /clone_lib="ZM_0.7-1.5_KB"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 977;
 Best Local Similarity 94.7%; Pred. No. 8.8e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 GACATGCCCGGCGCATGTC 20
 Db 831 GACATGCCCGGCGCATGTC 813

RESULT 23
 CL640722 179 bp mRNA linear GSS 30-JUN-2004
 LOCUS M003D05 GATC Gene Trap Library GV05C03 Mus musculus cDNA clone
 DEFINITION M003D05, mRNA sequence.
 ACCESSION CL640722
 VERSION CL640722.1 GI:49489169
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 179)
 Hansen, J., Flores, T., van Sloun, P., Fuchtbauer, E.M., Vauti, F.,
 Arnold, H.H., Schutgen, F., Wurst, W., Von Melchner, H. and Ruiz, P.
 A large-scale, gene-driven mutagenesis approach for the functional
 analysis of the mouse genome
 Proc. Natl. Acad. Sci. U.S.A. 100 (17), 9918-9922 (2003)
 22810117
 PUBMED 12904583
 COMMENT Contact: GGTC
 German Genetrap Consortium (GGTC)
 Email: info@genetrap.de
 Ubetageo gene trap. Sequence tag generated by 5'RACE. Additional
 sequence information can be found at:
 'http://genetrap.gsf.de/project/web_new/database/result_clone.html?target_id=M003D05' ES cell line harboring insertion mutation of
 'http://genetrap.gsf.de/project/web_new/order_clones/howtoorder.htm' 1' Inhouse Sequence Identifier: 01806
 Class: Gene Trap.

FEATURES
 source
 Location/Qualifiers
 1. 179
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129 Sv"
 /db_xref="taxon:10090"
 /clone="M003D05"
 /sex="Male"
 /cell_type="Embryonic stem cell"
 /cell_line="ES cells [129/Sv x 129X1 (formerly 129/Sv)]"
 /clone_lib="GGTC Gene Trap Library GV05C03"
 /note="Vector: Ubetageo"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 179;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 CY 1 GGACATGCCCGGCGCATGTC 20
 Db 96 GGGCTTCCCGGCGCATGTC 115

RESULT 24
 CL640722 179 bp mRNA linear GSS 30-JUN-2004
 LOCUS M003D05 GATC Gene Trap Library GV05C03 Mus musculus cDNA clone
 DEFINITION M003D05, mRNA sequence.
 ACCESSION CL640722
 VERSION CL640722.1 GI:49489169
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 179)
 Hansen, J., Flores, T., van Sloun, P., Fuchtbauer, E.M., Vauti, F.,
 Arnold, H.H., Schutgen, F., Wurst, W., Von Melchner, H. and Ruiz, P.
 A large-scale, gene-driven mutagenesis approach for the functional
 analysis of the mouse genome
 Proc. Natl. Acad. Sci. U.S.A. 100 (17), 9918-9922 (2003)
 22810117
 PUBMED 12904583
 COMMENT Contact: GGTC
 German Genetrap Consortium (GGTC)
 Email: info@genetrap.de
 Ubetageo gene trap. Sequence tag generated by 5'RACE. Additional
 sequence information can be found at:
 'http://genetrap.gsf.de/project/web_new/database/result_clone.html?target_id=M003D05' ES cell line harboring insertion mutation of
 'http://genetrap.gsf.de/project/web_new/order_clones/howtoorder.htm' 1' Inhouse Sequence Identifier: 01806
 Class: Gene Trap.

FEATURES
 source
 Location/Qualifiers
 1. 179
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129 Sv"
 /db_xref="taxon:10090"
 /clone="M003D05"
 /sex="Male"
 /cell_type="Embryonic stem cell"
 /cell_line="ES cells [129/Sv x 129X1 (formerly 129/Sv)]"
 /clone_lib="GGTC Gene Trap Library GV05C03"
 /note="Vector: Ubetageo"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 179;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCGCATGTC 20
 Db 115 GGACATGCCCGGCGCATGTC 96

RESULT 25
 B2691000 221 bp DNA linear GSS 14-FEB-2003
 LOCUS M003D05 GV05C03 Mus musculus genomic clone M003D05, genomic survey
 DEFINITION B2691000
 ACCESSION B2691000
 VERSION B2691000.1 GI:28382804
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 221)
 AUTHORS Hansen, J., Floss, T., van Sloun, P., Fuchtbauer, E.M., Vauti, F., Arnold, H.H., Schutgen, P., Wurst, W., Von Melchner, H. and Ruiz, P.
 TITLE A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (17), 9918-9922 (2003)
 MEDLINE 22810117
 PUBMED 12904583
 COMMENT Contact: Wurst W
 Institute of Developmental Genetics
 GSF - National Research Center for Environment and Health
 Ingolstaedter Landstrasse 1, D-85764 Neuherberg, Germany
 Tel: 49-89-3187-4110
 Fax: 49-89-3187-3099
 Email: wurst@gsf.de
 U3btageo gene trap. Tag generated by RACE. Additional sequence information can be found at:
 'http://genetrapp.gsf.de/project/web_new/database/result_clone.html?clone_id=M003D05'. ES cell line harboring insertion mutation of target gene is available at:
 'http://genetrapp.gsf.de/project/web_new/order_clones/howtoorder.htm'
 1
 Inhouse Sequence Identifier: 01806
 Class: Gene Trap.
 Location/Qualifiers
 1..221
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="129 Sv"
 /db_xref="taxon:10090"
 /clone="M003D05"
 /sex="Male"
 /cell_type="Embryonic stem cell"
 /cell_line="ES cells [129/Sv x 129X1 (formerly 129/Sv)]"
 /clone_lib="GV05C03"
 /note="Vector: U3btageo"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 8; Length 221;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCGCATGTCC 20
 |||||
 DB 65 GGACATGCCCGGCGCAAGCC 84

RESULT 26
 BZ691000/c 221 bp DNA linear GSS 14-FEB-2003
 LOCUS M003D05.GV05C03 Mus musculus genomic clone M003D05, genomic survey sequence.
 ACCESSION BZ691000
 VERSION BZ691000.1 GI:28382804
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 221)
 Hansen, J., Floss, T., van Sloun, P., Fuchtbauer, E.M., Vauti, F., Arnold, H.H., Schutgen, P., Wurst, W., Von Melchner, H. and Ruiz, P.
 TITLE A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (17), 9918-9922 (2003)
 MEDLINE 22810117
 PUBMED 12904583
 COMMENT Contact: Wurst W
 Institute of Developmental Genetics
 GSF - National Research Center for Environment and Health
 Ingolstaedter Landstrasse 1, D-85764 Neuherberg, Germany
 Tel: 49-89-3187-4110
 Fax: 49-89-3187-3099

REFERENCE Email: wurst@gsf.de
 U3btageo gene trap. Tag generated by RACE. Additional sequence information can be found at:
 'http://genetrapp.gsf.de/project/web_new/database/result_clone.html?clone_id=M003D05'. ES cell line harboring insertion mutation of target gene is available at:
 'http://genetrapp.gsf.de/project/web_new/order_clones/howtoorder.htm'
 1
 Inhouse Sequence Identifier: 01806
 Class: Gene Trap.
 Location/Qualifiers
 1..221
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="129 Sv"
 /db_xref="taxon:10090"
 /clone="M003D05"
 /sex="Male"
 /cell_type="Embryonic stem cell"
 /cell_line="ES cells [129/Sv x 129X1 (formerly 129/Sv)]"
 /clone_lib="GV05C03"
 /note="Vector: U3btageo"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 8; Length 221;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GGACATGCCCGGCGCATGTCC 20
 |||||
 DB 84 GGCGTTCGCGGCGCATGTCC 65

RESULT 27
 AV322349 227 bp mRNA linear EST 09-NOV-1999
 LOCUS AV322349
 DEFINITION AV322349 RIKEN full-length enriched, 14 days embryo thymus Mus musculus cDNA clone 6130401D12 3' similar to D86726 Mouse mRNA for mMS5, mRNA sequence.
 ACCESSION AV322349
 VERSION AV322349.1 GI:6292237
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 227)
 Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kuwahara, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomioka, N., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Kono, H., et al. 1999)
 Unpublished (1999)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Shohiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsf.riken.jp, URL: http://genome.gsc.riken.jp/
 Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsunaga, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 TITLE Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 JOURNAL Itoh, M., Kikunishi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y. Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303 19-44 (1999)

FEATURES

Source

```

Location/Qualifiers
1..227
/organism="Mus musculus"
/mol_type="RNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="6130401D12"
/sex="mixed"
/tissue_type="thymus"
/dev_stage="14 days embryo"
/lab_host="DH1.0B"
/clone_lib="RIKEN full-length enriched, 14 days embryo
cDNA library"

```

ORIGIN

Query Match	84.0%;	Score 16.8;	DB 1;	Length 227;
Best Local Similarity	90.0%;	Pred. No. 1.6e+03;		
Matches 18;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

```

QY      1  GGACATGCCCGGGCATGTCC  20
          ||||| ||||| ||||| |||||
Db      35  GGACACGCCCGGGCATTTCC  54

```

RESULT	28
AV322349/c	
LOCUS	
DEFINITION	AV322349 RIKEN full-length enriched, 14 days embryo thymus Mus
	227 bp mRNA linear EST 09-NOV-1996

ACCESSION	AV322349
VERSION	AV322349.1
	GI:6292237

SOURCE ORGANISM	Mus musculus (house mouse) Mus musculus
-----------------	--

REFERENCE AUTHORS

Komuro, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Koizumi, Y., Koya, S., Kusabeke, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sanjo, M., Sakato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugihara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateono, M., Tomimaga, N., Tsunoda, Y., Watanishi, A., Watanabe, S., Yamamura, T., Yasunishi, A.

TITLE
JOURNAL
COMMENT

Yokota, T., Yoshida, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al. 1999)
Unpublished (1999)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 95 (7), 3455-3460 (1998)

Itch, M., Kikunishi, T., Akiyama, J., Shibata, K., Iwano, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, T., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

Source

```

location/Qualifiers
1. 227
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="6130401D12"
/bex="mixed"
/tissue_type="thymus"
/dev_stage="14 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 14 days embryo"
thymus"
thymus"

```

ORIGIN

Query Match	84.0%;	Score 16.8;	DB 1;	Length 227;
Best Local Similarity	90.0%;	Pred. No. 1.6e+03;		
Matches 18; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

2y 1 GGACATGCCCGGGCATGTCC 20
 ||| ||||| ||||| |||||
 Db 54 GGAATGCCCGGGCGTGTCC 35

RESULT 29				
AZ377623	AZ377623	269 bp	DNA	linear
LOCUS				
DEFINITION	IM0132L01F Mouse 10kb plasmid UUGC1M library Mms musculus genomic			
ACCESSION	clone UUGC1M0132L01 F, genomic survey sequence.			
	AZ377623			

VERSION AZ377623.1 GI:10491323
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 269)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weis,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0132 row: L column: 01
 Seq primer: CGTGTAAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 269.
 Location/Qualifiers
 1. 269
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUC1M0132L01"
 /sex="Male"
 /lab_host="B. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_1lb="Mouse 10kb plasmid UUC1M library"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent B. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 269;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 1 GGACATGCCCGGCGCATGTCC 20
 164 GGCGCTGCCCGGCGCATGTCC 183

RESULT 30
 AZ377623/C
 LOCUS 269 bp DNA linear GSS 02-OCT-2000
 DEFINITION IM0132L01P Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC1M0132L01 F, genomic survey sequence.
 ACCESSION AZ377623

VERSION AZ377623.1 GI:10491323
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 269)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weis,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0132 row: L column: 01
 Seq primer: CGTGTAAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 269.
 Location/Qualifiers
 1. 269
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUC1M0132L01"
 /sex="Male"
 /lab_host="B. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_1lb="Mouse 10kb plasmid UUC1M library"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent B. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 269;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 1 GGACATGCCCGGCGCATGTCC 20
 183 GGACATGCCCGGCGCATGTCC 164

RESULT 31
 AZ505789
 LOCUS 274 bp DNA linear GSS 05-OCT-2000
 DEFINITION IM0346E12R Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC1M0346E12 R, genomic survey sequence.
 ACCESSION AZ505789

VERSION AZ505789.1 GI:10687105
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus. 1 (bases 1 to 274)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0346 row: E column: 12
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 274.
location/Qualifiers
1..274
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCGCM0346E12"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUCGCM library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 84.0%; Score 16.8; DB 8; Length 274;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 GGACATGCCCGGCGCATGTCC 20
DB 99 GGGCTGCGCCGGCGCATGTCC 118

RESULT 32
AZ505789/c 274 bp DNA linear GSS 05-OCT-2000
LOCUS 1M0346E12R Mouse 10kb plasmid UUCGCM library Mus musculus genomic
DEFINITION clone UUCGCM0346E12 R, genomic survey sequence.
ACCESSION AZ505789

VERSION AZ505789.1 GI:10687105
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus. 1 (bases 1 to 274)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0346 row: E column: 12
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 274.
location/Qualifiers
1..274
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCGCM0346E12"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUCGCM library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 84.0%; Score 16.8; DB 8; Length 274;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 GGACATGCCCGGCGCATGTCC 20
DB 118 GGACATGCCCGGCGCATGTCC 99

RESULT 33
BP090312 280 bp mRNA linear EST 30-JUN-2004
LOCUS BP090312 Chlamydomonas reinhardtii C9 various conditions
DEFINITION Chlamydomonas reinhardtii cDNA clone MX22h10_r 5', mRNA sequence.
ACCESSION BP090312

VERSION BP090312.1 GI:49462399
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii
 ORGANISM Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadaceae; Chlamydomonas.
 REFERENCE 1 (bases 1 to 280)
 AUTHORS Asamizu,E., Nakamura,Y., Mura,K., Fukuzawa,H., Fujiwara,S.,
 Hitono,M., Iwamoto,K., Matsumoto,Y., Minagawa,J., Shimogawara,K.,
 Takahashi,Y. and Tabata,S.
 TITLE Establishment of Publicly Available cDNA Material and Information
 Resource of Chlamydomonas reinhardtii (Chlorophyta), to Facilitate
 Gene Function Analysis
 JOURNAL Phycologia (2004) In press
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 FEATURES
 source
 Location/Qualifiers
 1..280
 /organism="Chlamydomonas reinhardtii"
 /mol_type="mRNA"
 /strain="C9"
 /db_xref="taxon:3055"
 ORIGIN
 Query Match 84.0%; Score 16.8; DB 5; Length 280;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GGACATGCCCGGCATGTC 20
 |||||
 187 GGATGTGCCCGGCATGTC 206
 Db
 RESULT 34 280 bp mRNA linear EST 30-JUN-2004
 BP090312/c Chlamydomonas reinhardtii C9 various conditions
 LOCUS Chlamydomonas reinhardtii cDNA clone WX21h10_r5', mRNA sequence.
 ACCESSION BP090312
 VERSION BP090312.1 GI:49462399
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii
 ORGANISM Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadaceae; Chlamydomonas.
 REFERENCE 1 (bases 1 to 280)
 AUTHORS Asamizu,E., Nakamura,Y., Mura,K., Fukuzawa,H., Fujiwara,S.,
 Hitono,M., Iwamoto,K., Matsumoto,Y., Minagawa,J., Shimogawara,K.,
 Takahashi,Y. and Tabata,S.
 TITLE Establishment of Publicly Available cDNA Material and Information
 Resource of Chlamydomonas reinhardtii (Chlorophyta), to Facilitate
 Gene Function Analysis
 JOURNAL Phycologia (2004) In press
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 FEATURES
 source
 Location/Qualifiers
 1..280
 /organism="Chlamydomonas reinhardtii"
 /mol_type="mRNA"
 /strain="C9"
 /db_xref="taxon:3055"

/clone="WX21h10_r"
 /clone_id="Chlamydomonas reinhardtii C9 various
 conditions"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; The cDNA library was made from a mixture of cells
 grown under various conditions"
 ORIGIN
 Query Match 84.0%; Score 16.8; DB 5; Length 280;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GGACATGCCCGGCATGTC 20
 |||||
 206 GGACATGCCCGGCATGTC 187
 Db
 RESULT 35 282 bp mRNA linear EST 23-JUN-2000
 BB093079 RIKEN full-length enriched, 12 days embryo, embryonic body
 LOCUS between diaphragm region and neck Mus musculus cDNA clone
 DEFINITION 9430041B07.3', mRNA Sequence.
 ACCESSION BB093079
 VERSION BB093079.1 GI:8676326
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mammalia; Euteleostomi; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 282)
 AUTHORS Komuro,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,
 Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
 Kiyosawa,M., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
 Kuwabake,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
 Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
 Shibata,K., Shibata,Y., Shigemoto,Y., Shingawa,A., Shiraki,T.,
 Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,
 Takahashi,F., Tomimaga,N., Toya,T., Tsunoda,Y., Watanishi,A.,
 Watanabe,S., Yamamura,T., Yamanaka,T., Yano,R., Yasunishi,A.,
 Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and
 Hayashizaki,Y.
 TITLE RIKEN Mouse ESTs (Komuro,H., et al.)
 JOURNAL Unpublished (2000)
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagoka,S.,
 Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Thermotabilization and thermocycling of thermostable enzymes by
 chitinase and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh,M., Katsunari,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
 Okazaki,Y. and Hayashizaki,Y.
 TITLE Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.
 FEATURES
 source
 Location/Qualifiers
 1..282
 /organism="Mus musculus"
 /mol_type="mRNA"


```

QY      1 GGACATGCCCGGGCATGTCC 20
          |||||
Db      36 GGACATGCCCGGGCAAGCCC 55

```

```

/db xref="taxon:10930"
/clone="9630041E07"
/tissue type="embryonic body between diaphragm region and neck"
/dev stage="12 days embryo"
/lab host="DH103"
/clone.lib="Riken full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck"
/note="Site_1: Salt; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGACAGAGAAGTCACAAGCGCTTTTTTTTTTTTAA 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGGACAGACTTCGAGTGTAATTAATTATCCCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI."

```

ACCESSION	BB093079	GI:8676326
VERSION	BB093079.1	
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

TITLE
JOURNAL
COMMENT

1 (bases 1 to 282)

Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hiraoane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Iizawa, M., Kadoro, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, R., Shibata, K., Shibata, Y., Shigenoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toyota, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaoka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yamana, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)

Unpublished (2000)

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Source

Thermosensitization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2), 520-524 (1998)

Itou, M., Kitsunai, T., Aitayama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

```

1. .282
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="g430041E07"
/tissue_type="embryonic body between diaphragm region and
neck"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_lib="Riken full-length enriched, 12 days embryo,
embryonic body between diaphragm region and neck"
/notes="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGGAGGAGGAGGATTCGAGGAGCTCTTTTCTTTTCTTTTAA 3'], cDNA was
prepared by using triethanolse thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 370.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGGAGGAGGATTCGAGGATTAATTAATTCCTCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI."

```

Query March	84.0%;	Score 16.8;	DB 2;	Length 282;
Best Local Similarity	90.0%;	Pred. No.1.6e+03;		
Matches 18;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

```

QY      1  GGACATGCCCGGGCATGTCC  20
          ||| ||||| ||||| |||||
Db      55  GGGCTTGCCCGGGCATGTCC  36

```

RESULT 37	
BBI57894	
LOCUS	283 bp mRNA linear EST 29-JUN-2000
DEFINITION	BBI57894 RIKEN full-length enriched, 16 days neonate thymus Mus
	musculus cDNA clone A130038G02 3', mRNA sequence.
ACCESSION	BBI57894
VERSION	BBI57894.1 GI:8813824
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCES

Carrinetti, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.
 Hirazawa, T., Horii, F., Ishii, Y., Ishikawa, T., Itoh, M.,
 Izawa, M., Kadota, K., Kawagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
 Kiyosawa, H., Kojima, Y., Komoto, S., Koye, S., Kurihara, C.,
 Kuwahabe, M., Matsuyama, T., Mikir, R., Mizuno, S., Nakamura, M., Oda, H.,
 Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
 Shibata, K., Shibata, Y., Shigemoto, Y., Shimaogawa, A., Shiraki, T.,
 Sugabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,

ORIGIN FLC I."

Query Match 84.0%; Score 16.8; DB 2; Length 283;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
 |||
 57 GGACCTTGCCCGGCGCATGTCC 38

Db

RESULT 39
 BB755436 283 bp mRNA linear EST 17-OCT-2001
 BB755436 RIKEN full-length enriched, melanocyte Mus musculus cDNA
 clone G270059B17 3', mRNA sequence.

DEFINITION
 BB755436
 BB755436.1 GI:16201995

ACCESSION
 BB755436.1 GI:16201995

VERSION
 BB755436.1 GI:16201995

KEYWORDS
 EST.

SOURCE
 Mus musculus (house mouse)

ORGANISM
 Mus musculus

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 283)

AUTHORS
 Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
 Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K.,
 Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Komoto,H., Kouda,M.,
 Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Nunasaki,R.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
 Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
 Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.

TITLE
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
 2001)

JOURNAL
 Unpublished (2001)

COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.jp, URL:http://genome.gsc.riken.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Matsura,S., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Watanishi,M., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A.
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Komoto,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.

FEATURES
 source
 1. 283
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G270059B17"
 /cell_type="melanocyte"

ORIGIN /clone_lib="RIKEN full-length enriched, melanocyte"

Query Match 84.0%; Score 16.8; DB 2; Length 283;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
 |||
 39 GGACATGCCCGGCGCATGTCC 58

Db

RESULT 40
 BB755436 283 bp mRNA linear EST 17-OCT-2001
 BB755436 RIKEN full-length enriched, melanocyte Mus musculus cDNA
 clone G270059B17 3', mRNA sequence.

DEFINITION
 BB755436
 BB755436.1 GI:16201995

ACCESSION
 BB755436.1 GI:16201995

VERSION
 BB755436.1 GI:16201995

KEYWORDS
 EST.

SOURCE
 Mus musculus (house mouse)

ORGANISM
 Mus musculus

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 283)

AUTHORS
 Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
 Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K.,
 Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Komoto,H., Kouda,M.,
 Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Nunasaki,R.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
 Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
 Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.

TITLE
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
 2001)

JOURNAL
 Unpublished (2001)

COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.jp, URL:http://genome.gsc.riken.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Matsura,S., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Watanishi,M., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A.
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Komoto,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.

FEATURES
 source
 1. 283
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G270059B17"
 /cell_type="melanocyte"

ORIGIN /clone_lib="RIKEN full-length enriched, melanocyte"

Query Match 84.0%; Score 16.8; DB 2; Length 283;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
 58 GGGCTTGGCCCGGCGCATGTCC 39

RESULT 41
 BY469224 287 bp mRNA linear EST 02-JAN-2003
 LOCUS BY469224 RIKEN full-length enriched, melanocyte Mus musculus cDNA
 DEFINITION clone G270064G22 3', mRNA sequence.
 ACCESSION BY469224
 VERSION BY469224.1 GI:26803603
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 287)
 AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nixaido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Haegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsumura, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochia, C., Corbani, L.B., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Guenichon, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltats, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yangisawa, M., Yang, I., Yang, L., Yuan, Z., Zavoian, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, D., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Saeki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 22354683

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saitama-cho, Tsukuba-shi, Ibaraki, 305-8575, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/
 Aikawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imocani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numata, K., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

FEATURES source

ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 287;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTCC 20
 44 GGGCTTGGCCCGGCGCATGTCC 63

RESULT 42
 BY469224

LOCUS BY469224 RIKEN full-length enriched, melanocyte Mus musculus cDNA
 DEFINITION clone G270064G22 3', mRNA sequence.
 ACCESSION BY469224
 VERSION BY469224.1 GI:26803603
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 287)
 AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nixaido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Haegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsumura, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochia, C., Corbani, L.B., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Guenichon, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltats, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Cells were provided by Drs. William J Pavan, Stacie Loftus, and Denise Larson (Division of Intramural Research Genetic Disease Research Branch National Human Genome Research Institute, National Institutes of Health (NIH) Building: 49, Room 4A82 49 Convent Drive MSC 4472 Bethesda, Maryland U.S.A) whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
 Location/Qualifiers
 1..287
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6J"
 /db_xref="taxon:10090"
 /clone="G270064G22"
 /cell_type="melanocyte"
 /clone_lib="RIKEN full-length enriched, melanocyte"

Verardo, R., Wagner, L., Mahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yangisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carinini, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, J., Kawai, J., Akawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

1246851
22354683

Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Alawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Cells were provided by Drs. William J Pavan, Stacie Loftus, and Denise Larson (Division of Intramural Research Genetic Disease Research Branch National Human Genome Research Institute, National Institutes of Health (NIH) Building: 49, Room 4A82 49 Convent Drive MSC 4472 Bethesda, Maryland U.S.A) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
SOURCE

Location/Qualifiers
1. .287
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G27064G22"
/cell_type="melanocyte"
/clone_lib="RIKEN full-length enriched, melanocyte"

ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 287;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GGACATGCCGGGATGCC 20
|||
Db 63 GGAGCTGCCGGGATGCC 44

RESULT 43

CK337629
LOCUS
DEFINITION

CK337629 288 bp mRNA linear EST 22-DEC-2003
C0330A03-3 NIA Mouse Undifferentiated ES Cell cDNA library (Long)
Mus musculus cDNA clone NIA: C0330A03 IMAGE:30007682 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CK337629 GI:40293242
EST.
Mus musculus (house mouse)

REFERENCE
AUTHORS
TITLE

Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method
Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL
MEDLINE
PUBMED
COMMENT

Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: C0330 row: A column: 03
Seq primer: -21M13 Forward
High quality sequence stop: 288
POLYA=Yes.

FEATURES
SOURCE

Location/Qualifiers
1. .288
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv x 129/Sv-CP"
/db_xref="niaEST:C0330A03-3"
/db_xref="taxon:10090"
/clone="NIA:C0330A03 IMAGE:30007682"
/tissue_type="Undifferentiated ES Cell"
/cell_line="R1 ES cells"
/lab_host="DH10B"
/lab_lib="NIA Mouse Undifferentiated ES Cell cDNA library (Long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://igsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). PMID: 11544199). Total RNAs were obtained from Dr. Kenneth R. Boheler (National Institute on Aging, USA). ES cells were cultured without feeder cells in the presence of LIF and BRU-conditioned media. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):

5'-GACTAGTTCGATCGGACGGCGCCCTTTT-3' from 14.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker ltr-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.4 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 288;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACATGCCCGGCGCATGTC 20
 Db 227 GGGCTTGCCCGGCGCATGTC 246
 RESULT 44
 CK337629 288 bp mRNA linear EST 22-DEC-2003
 LOCUS C0330A03-3 NIA Mouse Undifferentiated ES Cell cDNA Library (long)
 DEFINITION Mus musculus cDNA clone NIA:C0330A03 IMAGE:30007682 3', mRNA
 sequence.
 ACCESSION CK337629
 VERSION CK337629.1 GI:40293242
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Piao, Y., Ko, N.T., Lian, M.K. and Ko, M.S.H.
 1 (bases 1 to 288)
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 Genome Res. 11 (9), 1553-1558 (2001)
 JOURNAL 21429098
 MEDLINE 11544199
 PUBMED
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 Plate: C0330 row: A column: 03
 Seq primer: -21M13 Forward
 High quality sequence stop: 288
 POLY(A)=yes.
 FEATURES
 source
 1..288
 location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129/Sv x 129/Sv-Cp"
 /db_xref="taxon:C0330A03-3"
 /db_xref="taxon:10090"
 /clone="NIA:C0330A03 IMAGE:30007682"
 /tissue_type="Undifferentiated ES Cell"
 /cell_line="R1 ES cells"
 /lab_host="DH10B"
 /clone_1lb="NIA Mouse Undifferentiated ES Cell cDNA
 Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://1gsun.grc.nia.nih.gov/cDNA). This is a
 long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
 obtained from Dr. Kenneth R. Boheler (National Institute
 on Aging, USA). ES cells were cultured without feeder
 cells in the presence of LIF and BRL-conditioned media.
 Double-stranded cDNAs were synthesized with an Oligo(dT)
 primer (Invitrogen):
 5'-PGACTAGTTCGATCGCGAGCGCGCCCTTTT-3' from
 14.2 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Loner-linker L1-SalI, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Tag polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.
 The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.4 kb. The library was constructed
 by Yulan Piao (NIA)."

ORIGIN
 Query Match 84.0%; Score 16.8; DB 7; Length 288;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GGACATGCCCGGCGCATGTC 20
 Db 246 GGACATGCCCGGCGCATGTC 227
 RESULT 45
 CK340780 288 bp mRNA linear EST 22-DEC-2003
 LOCUS K0109H06-3 NIA Mouse Hematopoietic Stem Cell (lin-/c-Kit+/Sca-1-)
 DEFINITION cDNA library (long) Mus musculus cDNA clone NIA:K0109H06
 IMAGE:30040313 3', mRNA sequence.
 ACCESSION CK340780
 VERSION CK340780.1 GI:40296393
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Piao, Y., Ko, N.T., Lian, M.K. and Ko, M.S.H.
 1 (bases 1 to 288)
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 Genome Res. 11 (9), 1553-1558 (2001)
 JOURNAL 21429098
 MEDLINE 11544199
 PUBMED
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 Plate: K0109 row: H column: 06
 Seq primer: -21M13 Forward
 High quality sequence stop: 288
 POLY(A)=yes.
 FEATURES
 source
 1..288
 location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6NCR"
 /db_xref="taxon:K0109H06-3"
 /db_xref="taxon:10090"
 /clone="NIA:K0109H06 IMAGE:30040313"
 /tissue_type="Hematopoietic Stem Cell
 (lin-/c-Kit+/Sca-1-)"
 /dev_stage="Age approx.10 weeks old"
 /lab_host="DH10B"
 /clone_1lb="NIA Mouse Hematopoietic Stem Cell
 (lin-/c-Kit+/Sca-1-) cDNA Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://1gsun.grc.nia.nih.gov/cDNA). This is a
 long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
 obtained from Drs. Dennis Taub, Dan Longo (National
 Institute on Aging, USA), Jonathan Keller (National Cancer
 Institute, USA). Double-stranded cDNAs were synthesized
 with an Oligo(dT) primer (Invitrogen):
 5'-PGACTAGTTCGATCGCGAGCGCGCCCTTTT-3' from
 2.4 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Loner-linker L1-SalI, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Tag polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon

100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 288;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTC 20
|||
227 GGGCTTGCCCGGCGCATGTC 246

RESULT 46

CK340780 288 bp mRNA linear EST 22-DEC-2003
LOCUS K0109H06-3 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit+/Sca-1-)
DEFINITION CDNA library (Long) Mus musculus CDNA clone NIA:K0109H06
IMAGE:30040313 3', mRNA sequence.

ACCESSION CK340780 GI:40296393
VERSION CK340780.1 GI:40296393
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Piao, Y., Ko, N.-T., Lim, M.-K. and Ko, M.-S.H.
TITLE 1 (bases 1 to 288)
CONSTRUCTION of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098
PubMed 11544199

COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6620, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: K0109 row: H column: 06
Seq primer: -21M13 Forward
High quality sequence stop: 288
POLY-A: Yes

FEATURES

Source

1..288
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6NCR"
/db_xref="taxon:10090"
/db_xref="taxon:10090"
/clone="NIA:K0109H06 IMAGE:30040313"
/tissue_type="Hematopoietic Stem Cell
/lin-/c-Kit+/Sca-1-"
/dev_stage="Age approx. 10 weeks old"
/lab_host="MDH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
/lin-/c-Kit+/Sca-1-") CDNA library (Long)"
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
NotI; Mouse CDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://igsun.grc.nia.nih.gov/CDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). PMID: 11544199). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTAGATCGGACGGCGCCCTTTT-3'] from
2.4 ug of total RNA, treated with T4 DNA polymerase, and

ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 288;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACATGCCCGGCGCATGTC 20
|||
246 GGACATGCCCGGCGCATGTC 227

RESULT 47

BB060455 290 bp mRNA linear EST 25-JUN-2000
LOCUS BB060455 RIKEN full-length enriched, in vitro fertilized eggs Mus
DEFINITION musculus CDNA clone 7420450C07 3', mRNA sequence.
IMAGE:30040313 3', mRNA sequence.

ACCESSION BB060455.1 GI:8467603
VERSION BB060455.1 GI:8467603
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Kanno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,
Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,
Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T.,
Sogabe, Y., Sugahara, Y., Suzuki, H., Tagawa, A.,
Takahashi, F., Tomimaga, N., Toyota, T., Tsunoda, Y., Wachihi, A.,
Watanabe, S., Yamamura, T., Yamana, I., Yano, R., Yasunishi, A.,
Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.

TITLE RIKEN Mouse ESTs (Kanno, H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physiological and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.jp, url: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagakura, S.,
Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermolabile and thermostable restriction enzymes by
the use of a high-throughput cloning system for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
13-44 (1999)
Please visit our web site (http://genome.riken.go.jp) for

FEATURES

source

further details.

Location/Qualifiers

```
1..290
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420450C07"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, in vitro
fertilized eggs"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGATTCGAGTATTAATTAATATCCGCCGCCGCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified plasmid pUC19(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI"
```

ORIGIN

Query Match Best Local Similarity 84.0%; Score 16.8; DB 2; Length 290;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGAACATGCCCGGCGATGTC 20
233 GGAACATGCCCGGCGATGTC 252

RESULT 48
BB060455/c 290 bp mRNA linear EST 25-JUN-2000
DEFINITION BB060455 RIKEN full-length enriched, in vitro fertilized eggs Mus
musculus cDNA clone 7420450C07 3', mRNA sequence.
ACCESSION BB060455
VERSION BB060455.1 GI:8467603
KEYWORDS EST.

SOURCE

Mus musculus (house mouse)

REFERENCE 1 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 290)
Komori, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Hori, F., Ishi, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koyama, S., Kurahara, C.,
Kusakabe, M., Matsumura, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,
Okazaki, Y., Ono, T., Owa, C., Sakai, R., Saito, C., Sato, K.,
Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T.,
Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,
Takahashi, P., Tomimaga, N., Toya, T., Tsunoda, Y., Watanabe, S.,
Watanabe, S., Yamamura, T., Yamahata, I., Yano, R., Yasunishi, A.,
Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M., and
Hayashizaki, Y.

RIKEN Mouse ESTs (Komori, H., et al.)

Unpublished (2000)

CONTACT: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)

FEATURES

source

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagata, S.,
Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Thermotabilization and thermooxidation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (12), 520-524 (1998)
Itoh, M., Kikuchi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Oza, Y., Muramatsu, M.,
Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.ritc.riken.go.jp) for
further details.

Location/Qualifiers

```
1..290
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420450C07"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, in vitro
fertilized eggs"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGATTCGAGTATTAATTAATATCCGCCGCCGCC 3']. cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGATTCGAGTATTAATTAATATCCGCCGCCGCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified plasmid pUC19(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI"
```

ORIGIN

Query Match Best Local Similarity 84.0%; Score 16.8; DB 2; Length 290;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGAACATGCCCGGCGATGTC 20
252 GGAACATGCCCGGCGATGTC 233

RESULT 49
AA289954 293 bp mRNA linear EST 14-APR-1997
DEFINITION AA289954 RIKEN full-length enriched, in vitro fertilized eggs Mus
musculus cDNA clone 7420450C07 3', mRNA sequence.
ACCESSION AA289954
VERSION AA289954.1 GI:1936162
KEYWORDS EST.

Mus musculus (house mouse)

REFERENCE 1 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 293)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

AUTHORS

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 16:14:28 ; Search time 313.579 Seconds
(without alignments)
329.824 Million cell updates/sec

Title: US-09-578-453-1

Perfect score: 18

Sequence: 1 CGACTGTGATTCCTCCAT 18

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US05_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US04_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US03_PUBCOMB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US02_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US01_PUBCOMB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US00_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US05_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US04_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US03_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US02_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US01_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US00_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	100.0	1627	9	US-09-917-800A-1491
C 2	18	100.0	1627	16	US-10-368-934-568
C 3	16.4	91.1	390	9	US-09-974-300-3673
C 4	16.4	91.1	724	18	US-10-425-115-32562
C 5	16.4	91.1	1173	14	US-10-038-010-5
C 6	16.4	91.1	1173	15	US-10-267-384-214
C 7	16.4	91.1	1173	18	US-10-746-558-1
C 8	16.4	91.1	1208	15	US-10-392-113-32
C 9	15.4	85.6	441	16	US-10-424-599-102369
C 10	15.4	85.6	458	13	US-10-027-632-285372
C 11	15.4	85.6	458	15	US-10-027-632-285372
C 12	15.4	85.6	584	18	US-10-653-047-6563

C 13	15.4	85.6	604	18	US-10-425-115-57974	Sequence 57974, A
C 14	15.4	85.6	608	14	US-10-052-283-349	Sequence 349, App
C 15	15.4	85.6	616	18	US-10-425-115-180675	Sequence 180675, A
C 16	15.4	85.6	853	18	US-10-425-115-180677	Sequence 180677, A
C 17	15.4	85.6	881	18	US-10-425-115-178848	Sequence 178848, A
C 18	15.4	85.6	1017	18	US-10-425-115-180678	Sequence 180678, A
C 19	15.4	85.6	1510	9	US-09-731-872-8	Sequence 8, Appl1
C 20	15.4	85.6	1510	10	US-09-876-997-8	Sequence 8, Appl1
C 21	15.4	85.6	1523	9	US-09-731-872-47	Sequence 47, Appl1
C 22	15.4	85.6	1523	10	US-09-876-997-47	Sequence 47, Appl1
C 23	15.4	85.6	1535	9	US-09-935-390A-15	Sequence 15, Appl1
C 24	15.4	85.6	1755	9	US-09-745-763-188	Sequence 188, App
C 25	15.4	85.6	1892	9	US-09-925-298-221	Sequence 221, App
C 26	15.4	85.6	1892	14	US-10-102-806-221	Sequence 221, App
C 27	15.4	85.6	154504	17	US-10-322-696-67	Sequence 67, Appl1
C 28	15.4	85.6	200000	17	US-10-672-764A-33	Sequence 33, Appl1
C 29	15	83.3	18	14	US-10-131-786-2	Sequence 2, Appl1
C 30	15	83.3	1024	13	US-10-202-193-5	Sequence 5, Appl1
C 31	15	83.3	430442	18	US-10-417-375-128	Sequence 128, App
C 32	14.8	82.2	225	18	US-10-425-115-65084	Sequence 65084, A
C 33	14.8	82.2	281	11	US-09-922-293-137650	Sequence 137650, A
C 34	14.8	82.2	350	16	US-10-242-535A-30560	Sequence 30560, A
C 35	14.8	82.2	350	16	US-10-085-783A-30560	Sequence 30560, A
C 36	14.8	82.2	417	9	US-09-954-456-1075	Sequence 1075, App
C 37	14.8	82.2	417	9	US-09-954-456-1774	Sequence 1774, App
C 38	14.8	82.2	417	9	US-09-880-107-81	Sequence 81, Appl1
C 39	14.8	82.2	474	10	US-09-918-995-19674	Sequence 19674, A
C 40	14.8	82.2	578	13	US-10-027-632-137650	Sequence 137650, A
C 41	14.8	82.2	578	15	US-10-027-632-137650	Sequence 137650, A
C 42	14.8	82.2	600	18	US-10-425-115-47892	Sequence 47892, A
C 43	14.8	82.2	645	17	US-10-767-701-519	Sequence 519, App
C 44	14.8	82.2	846	16	US-10-425-115-108111	Sequence 108111, A
C 45	14.8	82.2	966	15	US-10-260-238-3388	Sequence 3388, App
C 46	14.8	82.2	966	15	US-10-238-075-579	Sequence 579, App
C 47	14.8	82.2	977	17	US-10-437-963-34632	Sequence 34632, A
C 48	14.8	82.2	1188	16	US-10-258-662-58	Sequence 58, Appl1
C 49	14.8	82.2	1448	16	US-10-424-599-82180	Sequence 82180, A
C 50	14.8	82.2	1826	17	US-10-437-963-35984	Sequence 35984, A
C 51	14.8	82.2	2498	15	US-10-085-959-44	Sequence 44, Appl1
C 52	14.8	82.2	2498	15	US-10-238-075-578	Sequence 578, App
C 53	14.8	82.2	2736	14	US-10-198-846-13456	Sequence 13456, A
C 54	14.8	82.2	2736	16	US-10-424-599-37574	Sequence 37574, A
C 55	14.8	82.2	5309	17	US-10-437-963-65649	Sequence 65649, A
C 56	14.8	82.2	5303	17	US-10-437-963-65652	Sequence 65652, A
C 57	14.8	82.2	653122	13	US-10-087-192-226	Sequence 226, App
C 58	14.8	82.2	1691139	14	US-10-067-515-1	Sequence 1, Appl1
C 59	14.8	82.2	1691139	16	US-10-419-723-1	Sequence 1, Appl1
C 60	14.4	80.0	19	9	US-09-969-373-851	Sequence 3851, App
C 61	14.4	80.0	35	16	US-10-223-560-9	Sequence 9, Appl1
C 62	14.4	80.0	123	9	US-09-969-373-1160	Sequence 1160, App
C 63	14.4	80.0	201	18	US-10-719-993-20905	Sequence 20905, A
C 64	14.4	80.0	201	18	US-10-719-993-50419	Sequence 50419, A
C 65	14.4	80.0	201	18	US-10-719-993-54062	Sequence 54062, A
C 66	14.4	80.0	331	9	US-09-864-761-17558	Sequence 17558, A
C 67	14.4	80.0	376	9	US-09-770-781-257	Sequence 257, App
C 68	14.4	80.0	413	10	US-09-918-995-8170	Sequence 8170, App
C 69	14.4	80.0	427	9	US-09-960-352-3398	Sequence 3398, App
C 70	14.4	80.0	433	10	US-09-918-995-8512	Sequence 8512, App
C 71	14.4	80.0	442	13	US-10-027-632-193264	Sequence 193264, A
C 72	14.4	80.0	442	15	US-10-027-632-193264	Sequence 193264, A
C 73	14.4	80.0	444	9	US-09-864-761-774	Sequence 774, App
C 74	14.4	80.0	466	13	US-10-027-632-83193	Sequence 83193, A
C 75	14.4	80.0	466	15	US-10-027-632-83193	Sequence 83193, A
C 76	14.4	80.0	477	10	US-09-918-995-9174	Sequence 9174, App
C 77	14.4	80.0	495	9	US-09-728-445-678	Sequence 678, App
C 78	14.4	80.0	517	13	US-10-027-632-87097	Sequence 87097, A
C 79	14.4	80.0	517	13	US-10-027-632-303688	Sequence 303688, A
C 80	14.4	80.0	517	15	US-10-027-632-87097	Sequence 87097, A
C 81	14.4	80.0	517	15	US-10-027-632-303688	Sequence 303688, A
C 82	14.4	80.0	563	16	US-10-424-599-35373	Sequence 35373, A
C 83	14.4	80.0	573	9	US-09-864-761-8280	Sequence 8280, App
C 84	14.4	80.0	598	15	US-10-029-813-6927	Sequence 6927, App
C 85	14.4	80.0	600	16	US-10-381-813-17	Sequence 17, Appl1

C 86	14.4	80.0	622	16	US-10-424-599-88609	Sequence 88609, A	C 159	13.8	76.7	50	18	US-10-667-141-28	Sequence 28, Appl
C 87	14.4	80.0	684	13	US-10-027-632-25713	Sequence 25713, A	C 160	13.8	76.7	60	10	US-09-908-975-17568	Sequence 17568, A
C 88	14.4	80.0	684	15	US-10-027-632-25713	Sequence 25713, A	C 161	13.8	76.7	200	9	US-09-864-761-22218	Sequence 22218, A
C 89	14.4	80.0	695	13	US-10-027-632-162409	Sequence 162409, A	C 162	13.8	76.7	201	17	US-10-741-601-25267	Sequence 25267, A
C 90	14.4	80.0	695	13	US-10-027-632-162410	Sequence 162410, A	C 163	13.8	76.7	201	18	US-10-719-993-15787	Sequence 15787, A
C 91	14.4	80.0	695	15	US-10-027-632-162409	Sequence 162409, A	C 164	13.8	76.7	201	18	US-10-719-993-25494	Sequence 25494, A
C 92	14.4	80.0	695	15	US-10-027-632-162410	Sequence 162410, A	C 165	13.8	76.7	201	18	US-10-719-993-28365	Sequence 28365, A
C 93	14.4	80.0	695	9	US-09-938-842A-2474	Sequence 2474, Ap	C 166	13.8	76.7	201	18	US-10-719-993-38413	Sequence 38413, A
C 94	14.4	80.0	864	11	US-09-938-842A-2474	Sequence 2474, Ap	C 167	13.8	76.7	201	18	US-10-719-993-38416	Sequence 38416, A
C 95	14.4	80.0	864	16	US-10-412-699B-219	Sequence 219, App	C 168	13.8	76.7	202	18	US-10-357-930-3881	Sequence 3881, Ap
C 96	14.4	80.0	920	9	US-09-770-445-419	Sequence 419, App	C 169	13.8	76.7	246	9	US-09-974-300-7851	Sequence 7851, Ap
C 97	14.4	80.0	993	17	US-10-437-963-85995	Sequence 85995, A	C 170	13.8	76.7	285	16	US-10-424-599-19280	Sequence 19280, A
C 98	14.4	80.0	1026	16	US-10-425-114-13022	Sequence 13022, A	C 171	13.8	76.7	300	9	US-09-294-093B-4843	Sequence 4843, Ap
C 99	14.4	80.0	1093	9	US-09-186-276B-53	Sequence 53, Appl	C 172	13.8	76.7	312	16	US-10-424-599-43175	Sequence 43175, A
C 100	14.4	80.0	1093	14	US-10-253-007-53	Sequence 53, Appl	C 173	13.8	76.7	320	15	US-10-264-283-47	Sequence 47, Appl
C 101	14.4	80.0	1093	14	US-10-253-007-53	Sequence 53, Appl	C 174	13.8	76.7	341	18	US-10-674-128A-5391	Sequence 5391, Ap
C 102	14.4	80.0	1098	9	US-09-738-626-2368	Sequence 8865, A	C 175	13.8	76.7	345	16	US-10-381-813-115	Sequence 115, App
C 103	14.4	80.0	1541	18	US-10-425-115-88465	Sequence 8865, A	C 176	13.8	76.7	345	16	US-10-381-813-115	Sequence 37196, A
C 104	14.4	80.0	1541	18	US-10-425-115-88465	Sequence 8865, A	C 177	13.8	76.7	348	18	US-10-425-115-37196	Sequence 22844, A
C 105	14.4	80.0	1850	17	US-10-433-485A-14	Sequence 32, Appl	C 178	13.8	76.7	375	16	US-10-242-535A-22844	Sequence 22844, A
C 106	14.4	80.0	1850	18	US-10-855-595-3	Sequence 3, Appl1	C 179	13.8	76.7	375	16	US-10-085-783A-22844	Sequence 55238, A
C 107	14.4	80.0	1850	18	US-10-855-532-3	Sequence 243, Appl	C 180	13.8	76.7	399	16	US-10-424-599-55238	Sequence 164935, A
C 108	14.4	80.0	2574	15	US-10-085-117-243	Sequence 219, App	C 181	13.8	76.7	394	18	US-10-425-115-164935	Sequence 34198, A
C 109	14.4	80.0	2774	15	US-10-278-536-219	Sequence 7715, Ap	C 182	13.8	76.7	400	17	US-10-767-795-2126	Sequence 2126, Ap
C 110	14.4	80.0	2822	18	US-10-723-860-7715	Sequence 808, App	C 183	13.8	76.7	405	18	US-10-357-930-34198	Sequence 43060, A
C 111	14.4	80.0	3417	15	US-10-104-047-808	Sequence 28, Appl	C 184	13.8	76.7	405	18	US-10-357-930-43060	Sequence 7901, Ap
C 112	14.4	80.0	4043	9	US-09-989-920-28	Sequence 98, App	C 185	13.8	76.7	419	18	US-10-357-930-7901	Sequence 35436, A
C 113	14.4	80.0	4967	9	US-09-834-975-558	Sequence 242, App	C 186	13.8	76.7	440	10	US-09-918-993-35436	Sequence 37849, A
C 114	14.4	80.0	4967	10	US-09-814-953-21662	Sequence 814, App	C 187	13.8	76.7	442	10	US-09-918-993-37849	Sequence 13050, A
C 115	14.4	80.0	5026	15	US-10-085-117-242	Sequence 51, Appl	C 188	13.8	76.7	443	18	US-10-357-930-13050	Sequence 136480, A
C 116	14.4	80.0	8674	14	US-10-044-090-814	Sequence 51, Appl	C 189	13.8	76.7	447	16	US-10-424-599-136480	Sequence 4730, Ap
C 117	14.4	80.0	8878	14	US-10-076-816-51	Sequence 512, Ap	C 190	13.8	76.7	452	9	US-09-867-701-4730	Sequence 22918, A
C 118	14.4	80.0	8878	14	US-10-017-724-5	Sequence 1512, Ap	C 191	13.8	76.7	459	16	US-10-282-122A-22918	Sequence 1254, Ap
C 119	14.4	80.0	32195	9	US-09-764-847-1512	Sequence 1512, Ap	C 192	13.8	76.7	459	17	US-09-960-352-1254	Sequence 63280, A
C 120	14.4	80.0	32195	14	US-10-092-154-1512	Sequence 7009, App	C 193	13.8	76.7	499	9	US-10-437-962-63280	Sequence 10744, Ap
C 121	14.4	80.0	42385	18	US-10-719-993-7009	Sequence 676, App	C 194	13.8	76.7	499	17	US-09-864-761-5648	Sequence 37552, A
C 122	14.4	80.0	114615	13	US-10-087-192-676	Sequence 241, App	C 195	13.8	76.7	474	9	US-09-864-761-10744	Sequence 235, App
C 123	14.4	80.0	165961	15	US-10-085-117-241	Sequence 2014, Ap	C 196	13.8	76.7	475	18	US-10-357-930-37852	Sequence 20859, A
C 124	14.4	80.0	248436	13	US-10-087-192-2014	Sequence 7065, Ap	C 197	13.8	76.7	486	15	US-10-091-007-235	Sequence 84363, A
C 125	14.4	80.0	283351	18	US-10-079-993-7065	Sequence 1, Appl1	C 198	13.8	76.7	498	17	US-10-437-962-84363	Sequence 169404, A
C 126	14.4	80.0	344805	18	US-10-779-271-1	Sequence 6815, Ap	C 199	13.8	76.7	499	18	US-10-425-115-169404	Sequence 12420, A
C 127	14.4	80.0	1580090	18	US-10-719-993-6615	Sequence 104931, A	C 200	13.8	76.7	501	9	US-09-864-761-12420	Sequence 264067, A
C 128	14.4	80.0	3309400	9	US-09-738-626-1	Sequence 80814, A	C 201	13.8	76.7	504	16	US-10-424-599-64067	Sequence 26440, A
C 129	14.4	77.8	413	16	US-10-424-599-104931	Sequence 80814, A	C 202	13.8	76.7	511	16	US-10-425-114-26440	Sequence 417, App
C 130	14.4	77.8	438	13	US-10-027-632-80814	Sequence 80814, A	C 203	13.8	76.7	514	10	US-09-551-621-417	Sequence 417, App
C 131	14.4	77.8	438	13	US-10-027-632-80815	Sequence 80815, A	C 204	13.8	76.7	514	9	US-09-604-287A-417	Sequence 417, App
C 132	14.4	77.8	438	15	US-10-027-632-80814	Sequence 80815, A	C 205	13.8	76.7	514	9	US-09-604-287A-417	Sequence 417, App
C 133	14.4	77.8	438	15	US-10-027-632-80815	Sequence 80815, A	C 206	13.8	76.7	514	13	US-09-551-621-417	Sequence 417, App
C 134	14.4	77.8	514	15	US-10-027-632-79374	Sequence 79374, A	C 207	13.8	76.7	514	14	US-10-076-622-417	Sequence 417, App
C 135	14.4	77.8	514	15	US-10-027-632-79374	Sequence 79374, A	C 208	13.8	76.7	514	15	US-10-076-622-417	Sequence 417, App
C 136	14.4	77.8	542	13	US-10-027-632-84022	Sequence 84022, A	C 209	13.8	76.7	514	15	US-10-124-805-417	Sequence 178845, A
C 137	14.4	77.8	542	15	US-10-027-632-84022	Sequence 226492, A	C 210	13.8	76.7	516	18	US-10-425-115-178845	Sequence 258, App
C 138	14.4	77.8	654	13	US-10-027-632-226492	Sequence 226492, A	C 211	13.8	76.7	519	9	US-09-604-287A-258	Sequence 258, App
C 139	14.4	77.8	654	13	US-10-027-632-226493	Sequence 226493, A	C 212	13.8	76.7	519	9	US-09-604-287A-258	Sequence 258, App
C 140	14.4	77.8	654	15	US-10-027-632-226493	Sequence 226493, A	C 213	13.8	76.7	519	9	US-09-339-338-258	Sequence 258, App
C 141	14.4	77.8	654	15	US-10-027-632-226493	Sequence 226493, A	C 214	13.8	76.7	519	10	US-09-551-621-258	Sequence 258, App
C 142	14.4	77.8	733	13	US-10-027-632-21556	Sequence 21556, A	C 215	13.8	76.7	519	13	US-10-007-805-258	Sequence 258, App
C 143	14.4	77.8	733	13	US-10-027-632-21556	Sequence 21556, A	C 216	13.8	76.7	519	13	US-10-007-805-258	Sequence 258, App
C 144	14.4	77.8	940	16	US-10-424-599-77765	Sequence 77765, A	C 217	13.8	76.7	519	15	US-10-124-805-258	Sequence 1527, A
C 145	14.4	77.8	1234	9	US-09-778-844-86	Sequence 86, Appl1	C 218	13.8	76.7	529	18	US-10-357-930-15927	Sequence 15876, A
C 146	14.4	77.8	1395	15	US-10-369-493-25339	Sequence 25439, A	C 219	13.8	76.7	539	18	US-10-425-115-132876	Sequence 149923, A
C 147	14.4	77.8	1395	16	US-10-793-639-167	Sequence 167, App	C 220	13.8	76.7	548	18	US-10-425-115-15923	Sequence 152, App
C 148	14.4	77.8	3438	16	US-10-062-674-1639	Sequence 1639, Ap	C 221	13.8	76.7	551	16	US-10-425-115-152	Sequence 21890, App
C 149	14.4	77.8	13524	18	US-10-719-993-6913	Sequence 6913, Ap	C 222	13.8	76.7	558	16	US-10-425-114-21890	Sequence 22055, A
C 150	14.4	77.8	13904	15	US-10-017-161-1977	Sequence 1625, Ap	C 223	13.8	76.7	561	16	US-10-425-114-22055	Sequence 169406, A
C 151	14.4	77.8	13904	15	US-10-992-798-1655	Sequence 31, Appl1	C 224	13.8	76.7	564	18	US-10-425-115-169406	Sequence 45755, A
C 152	14.4	77.8	39412	17	US-10-043-160-31	Sequence 5, Appl1	C 225	13.8	76.7	567	16	US-10-357-930-90678	Sequence 90678, A
C 153	14.4	77.8	255439	18	US-10-719-993-6799	Sequence 6799, Ap	C 226	13.8	76.7	568	18	US-10-793-032-4	Sequence 4, Appl1
C 154	14.4	77.8	255439	18	US-10-098-263B-115019	Sequence 309, App	C 227	13.8	76.7	572	13	US-10-027-632-22553	Sequence 22553, A
C 155	13.8	76.7	25	10	US-09-747-377-309	Sequence 309, App	C 228	13.8	76.7	572	13	US-10-027-632-22553	Sequence 128672, A
C 156	13.8	76.7	34	10	US-10-105-611-309	Sequence 27, Appl	C 229	13.8	76.7	575	13	US-10-027-632-128672	Sequence 128672, A
C 157	13.8	76.7	49	18	US-10-667-141-27		C 230	13.8	76.7	575	13	US-10-027-632-128672	
C 158	13.8	76.7	18	18	US-10-667-141-27		C 231	13.8	76.7	575	15	US-10-027-632-128672	

232	13.8	76.7	576	16	US-10-425-114-27205	Sequence 27205, A
C 233	13.8	76.7	582	17	US-10-021-323-7768	Sequence 7768, Ap
C 234	13.8	76.7	589	18	US-10-357-930-56981	Sequence 56981, A
C 235	13.8	76.7	598	13	US-10-027-632-230606	Sequence 230606, A
C 236	13.8	76.7	598	15	US-10-027-632-230606	Sequence 230606, A
C 237	13.8	76.7	601	13	US-10-027-632-24078	Sequence 24078, A
C 238	13.8	76.7	601	13	US-10-027-632-133294	Sequence 133294, A
C 239	13.8	76.7	601	13	US-10-027-632-133295	Sequence 133295, A
C 240	13.8	76.7	601	15	US-10-027-632-24078	Sequence 24078, A
C 241	13.8	76.7	601	15	US-10-027-632-133294	Sequence 133294, A
C 242	13.8	76.7	603	17	US-10-027-632-133295	Sequence 133295, A
C 243	13.8	76.7	606	18	US-10-425-115-154969	Sequence 65496, A
C 244	13.8	76.7	610	13	US-10-027-632-257253	Sequence 257253, A
C 245	13.8	76.7	610	13	US-10-027-632-257253	Sequence 257253, A
C 246	13.8	76.7	612	15	US-10-027-632-257253	Sequence 257253, A
C 247	13.8	76.7	621	17	US-10-437-963-84362	Sequence 84362, A
C 248	13.8	76.7	621	13	US-10-027-632-185178	Sequence 185178, A
C 249	13.8	76.7	627	15	US-10-027-632-185178	Sequence 185178, A
C 250	13.8	76.7	642	13	US-10-027-632-48134	Sequence 48134, A
C 251	13.8	76.7	642	13	US-10-425-115-178847	Sequence 178847, A
C 252	13.8	76.7	647	18	US-10-027-632-105460	Sequence 105460, A
C 253	13.8	76.7	652	13	US-10-027-632-105460	Sequence 105460, A
C 254	13.8	76.7	652	15	US-10-437-963-100249	Sequence 100249, A
C 255	13.8	76.7	652	15	US-10-027-632-105460	Sequence 105460, A
C 256	13.8	76.7	661	17	US-10-027-632-243275	Sequence 243275, A
C 257	13.8	76.7	677	13	US-10-027-632-243275	Sequence 243275, A
C 258	13.8	76.7	682	13	US-10-027-632-243275	Sequence 243275, A
C 259	13.8	76.7	682	15	US-10-027-632-22630	Sequence 22630, A
C 260	13.8	76.7	684	16	US-10-424-599-135530	Sequence 135530, A
C 261	13.8	76.7	687	18	US-10-357-930-22538	Sequence 22538, A
C 262	13.8	76.7	687	18	US-10-357-930-22538	Sequence 22538, A
C 263	13.8	76.7	696	18	US-10-425-115-148874	Sequence 148874, A
C 264	13.8	76.7	700	9	US-09-920-300A-1683	Sequence 1683, Ap
C 265	13.8	76.7	700	15	US-10-033-528-1683	Sequence 1683, Ap
C 266	13.8	76.7	700	15	US-10-033-528-1683	Sequence 1683, Ap
C 267	13.8	76.7	708	18	US-10-027-632-101426	Sequence 101426, A
C 268	13.8	76.7	708	18	US-10-027-632-101426	Sequence 101426, A
C 269	13.8	76.7	709	16	US-10-424-599-69223	Sequence 69223, A
C 270	13.8	76.7	714	13	US-10-027-632-16482	Sequence 16482, A
C 271	13.8	76.7	714	15	US-10-027-632-16482	Sequence 16482, A
C 272	13.8	76.7	735	17	US-10-767-701-14807	Sequence 14807, A
C 273	13.8	76.7	735	13	US-10-027-632-135508	Sequence 135508, A
C 274	13.8	76.7	735	15	US-10-027-632-135508	Sequence 135508, A
C 275	13.8	76.7	737	18	US-10-425-115-172628	Sequence 172628, A
C 276	13.8	76.7	739	13	US-10-027-632-101426	Sequence 101426, A
C 277	13.8	76.7	739	15	US-10-027-632-101426	Sequence 101426, A
C 278	13.8	76.7	748	17	US-10-437-963-55955	Sequence 55955, A
C 279	13.8	76.7	754	18	US-10-425-115-64349	Sequence 64349, A
C 280	13.8	76.7	754	18	US-10-425-115-64349	Sequence 64349, A
C 281	13.8	76.7	758	13	US-10-027-632-110325	Sequence 110325, A
C 282	13.8	76.7	758	15	US-10-027-632-110325	Sequence 110325, A
C 283	13.8	76.7	759	13	US-10-027-632-29911	Sequence 29911, A
C 284	13.8	76.7	759	15	US-10-027-632-29911	Sequence 29911, A
C 285	13.8	76.7	765	9	US-09-912-787-11	Sequence 11, Appl
C 286	13.8	76.7	765	9	US-09-912-787-70	Sequence 70, Appl
C 287	13.8	76.7	765	13	US-10-014-326-56	Sequence 56, Appl
C 288	13.8	76.7	765	13	US-10-014-326-56	Sequence 56, Appl
C 289	13.8	76.7	766	18	US-10-425-115-117779	Sequence 117779, A
C 290	13.8	76.7	781	18	US-10-425-115-114438	Sequence 114438, A
C 291	13.8	76.7	782	17	US-10-767-701-14806	Sequence 14806, A
C 292	13.8	76.7	792	17	US-09-938-842A-1950	Sequence 1950, Ap
C 293	13.8	76.7	792	11	US-09-938-842A-1950	Sequence 1950, Ap
C 294	13.8	76.7	792	11	US-10-425-114-35743	Sequence 35743, A
C 295	13.8	76.7	859	9	US-09-764-877-971	Sequence 971, Appl
C 296	13.8	76.7	859	16	US-10-242-515-971	Sequence 971, Appl
C 297	13.8	76.7	880	18	US-10-425-115-135103	Sequence 135103, A
C 298	13.8	76.7	887	18	US-10-425-115-147019	Sequence 147019, A
C 299	13.8	76.7	891	17	US-10-437-963-94739	Sequence 94739, A
C 300	13.8	76.7	900	17	US-10-437-963-94628	Sequence 94628, A
C 301	13.8	76.7	905	9	US-09-954-106-1198	Sequence 1198, Ap
C 302	13.8	76.7	905	9	US-09-880-107-424	Sequence 424, Appl
C 303	13.8	76.7	915	15	US-10-369-493-33793	Sequence 33793, A
C 304	13.8	76.7	946	18	US-10-425-115-147008	Sequence 147008, A
C 305	13.8	76.7	967	15	US-10-106-698-760	Sequence 760, Appl
C 306	13.8	76.7	981	16	US-10-424-599-74187	Sequence 74187, A
C 307	13.8	76.7	990	17	US-10-437-963-99293	Sequence 99293, A
C 308	13.8	76.7	999	16	US-10-425-114-25917	Sequence 25917, A
C 309	13.8	76.7	1016	9	US-09-823-245A-451	Sequence 451, Appl
C 310	13.8	76.7	1018	17	US-10-767-701-5775	Sequence 5775, Ap
C 311	13.8	76.7	1021	16	US-10-321-039-11	Sequence 17, Appl
C 312	13.8	76.7	1022	15	US-10-074-511-58	Sequence 51, Appl
C 313	13.8	76.7	1035	16	US-10-425-114-22420	Sequence 22420, A
C 314	13.8	76.7	1041	15	US-09-371-307-79	Sequence 79, Appl
C 315	13.8	76.7	1041	15	US-10-401-321-79	Sequence 79, Appl
C 316	13.8	76.7	1046	9	US-09-925-297-895	Sequence 295, Appl
C 317	13.8	76.7	1047	16	US-10-425-114-22427	Sequence 22427, A
C 318	13.8	76.7	1053	9	US-09-770-445-157	Sequence 157, Appl
C 319	13.8	76.7	1059	16	US-10-425-114-7086	Sequence 7086, Ap
C 320	13.8	76.7	1059	16	US-10-425-114-7086	Sequence 7086, Ap
C 321	13.8	76.7	1130	18	US-10-425-115-149921	Sequence 149921, A
C 322	13.8	76.7	1230	16	US-10-282-122A-41808	Sequence 41808, A
C 323	13.8	76.7	1238	18	US-10-425-115-42420	Sequence 42420, A
C 324	13.8	76.7	1272	16	US-10-282-122A-35694	Sequence 35694, A
C 325	13.8	76.7	1302	16	US-10-424-599-17155	Sequence 17155, A
C 326	13.8	76.7	1305	16	US-10-282-122A-20074	Sequence 20074, A
C 327	13.8	76.7	1305	17	US-10-437-963-19573	Sequence 19573, A
C 328	13.8	76.7	1316	18	US-10-425-115-135105	Sequence 135105, A
C 329	13.8	76.7	1373	9	US-09-263-959-285	Sequence 285, Appl
C 330	13.8	76.7	1395	9	US-09-938-842A-724	Sequence 724, Appl
C 331	13.8	76.7	1395	11	US-09-938-842A-724	Sequence 724, Appl
C 332	13.8	76.7	1431	10	US-09-992-600A-87	Sequence 87, Appl
C 333	13.8	76.7	1431	10	US-09-992-600A-87	Sequence 87, Appl
C 334	13.8	76.7	1431	10	US-09-924-340-89	Sequence 89, Appl
C 335	13.8	76.7	1431	10	US-09-924-340-89	Sequence 89, Appl
C 336	13.8	76.7	1431	10	US-09-992-095E-87	Sequence 87, Appl
C 337	13.8	76.7	1431	10	US-09-992-095E-89	Sequence 89, Appl
C 338	13.8	76.7	1431	10	US-09-999-570-87	Sequence 87, Appl
C 339	13.8	76.7	1431	10	US-09-999-570-87	Sequence 87, Appl
C 340	13.8	76.7	1431	14	US-10-000-489-87	Sequence 89, Appl
C 341	13.8	76.7	1431	14	US-10-000-489-89	Sequence 89, Appl
C 342	13.8	76.7	1431	14	US-10-000-986-87	Sequence 87, Appl
C 343	13.8	76.7	1431	14	US-10-000-986-89	Sequence 89, Appl
C 344	13.8	76.7	1431	15	US-10-154-678-87	Sequence 87, Appl
C 345	13.8	76.7	1431	15	US-10-154-678-89	Sequence 89, Appl
C 346	13.8	76.7	1431	15	US-10-001-142-87	Sequence 87, Appl
C 347	13.8	76.7	1431	15	US-10-001-142-89	Sequence 89, Appl
C 348	13.8	76.7	1433	18	US-10-425-115-52898	Sequence 52898, A
C 349	13.8	76.7	1506	9	US-09-886-241-4	Sequence 4, Appl1
C 350	13.8	76.7	1513	9	US-09-920-300A-1682	Sequence 1682, Ap
C 351	13.8	76.7	1513	13	US-10-033-528-1682	Sequence 1682, Ap
C 352	13.8	76.7	1513	15	US-10-205-823-116	Sequence 116, Appl
C 353	13.8	76.7	1513	15	US-10-177-193-153	Sequence 153, Appl
C 354	13.8	76.7	1513	15	US-10-099-926-1682	Sequence 1682, Appl
C 355	13.8	76.7	1518	17	US-10-738-455-2	Sequence 2, Appl1
C 356	13.8	76.7	1527	18	US-10-425-115-150275	Sequence 150275, A
C 357	13.8	76.7	1551	17	US-10-437-963-437	Sequence 437, Ap
C 358	13.8	76.7	1567	18	US-10-425-115-9029	Sequence 9029, Ap
C 359	13.8	76.7	1668	18	US-10-739-930-439	Sequence 439, Appl
C 360	13.8	76.7	1677	13	US-10-044-090-140	Sequence 140, Appl
C 361	13.8	76.7	1718	18	US-10-425-115-29698	Sequence 29698, A
C 362	13.8	76.7	1781	13	US-10-027-632-200093	Sequence 200093, A
C 363	13.8	76.7	1800	15	US-10-027-632-200093	Sequence 200093, A
C 364	13.8	76.7	1800	18	US-10-739-930-89	Sequence 89, Appl
C 365	13.8	76.7	1841	16	US-10-435-696-69	Sequence 69, Appl
C 366	13.8	76.7	1853	9	US-09-745-763-84	Sequence 84, Appl
C 367	13.8	76.7	1893	16	US-10-424-599-42110	Sequence 42110, A
C 368	13.8	76.7	2002	9	US-09-925-301-330	Sequence 230, Appl
C 369	13.8	76.7	2020	9	US-09-984-245-49	Sequence 49, Appl
C 370	13.8	76.7	2020	10	US-09-966-262-49	Sequence 49, Appl
C 371	13.8	76.7	2020	10	US-09-983-395-49	Sequence 49, Appl
C 372	13.8	76.7	2020	14	US-10-059-395-49	Sequence 49, Appl
C 373	13.8	76.7	2020	14	US-10-143-090-49	Sequence 49, Appl
C 374	13.8	76.7	2067	16	US-10-282-122A-26924	Sequence 26924, A
C 375	13.8	76.7	2129	17	US-10-437-963-26460	Sequence 26460, A
C 376	13.8	76.7	2179	17	US-10-115-635-117	Sequence 117, Appl
C 377	13.8	76.7	2199	15	US-10-205-194-56	Sequence 56, Appl

C 378	13.8	76.7	2199	16	US-10-388-934-636	Sequence 636, App
C 379	13.8	76.7	2238	15	US-10-260-238-529	Sequence 529, App
C 380	13.8	76.7	2244	17	US-10-437-963-13344	Sequence 13344, A
C 381	13.8	76.7	2207	15	US-10-120-988-314	Sequence 314, App
C 382	13.8	76.7	2340	17	US-10-437-963-65515	Sequence 65515, A
C 383	13.8	76.7	2420	18	US-10-357-930-25066	Sequence 25066, A
C 384	13.8	76.7	2455	16	US-10-424-599-124227	Sequence 124227, A
C 385	13.8	76.7	2510	16	US-10-108-260A-1354	Sequence 1354, App
C 386	13.8	76.7	2518	13	US-10-027-632-103157	Sequence 103157, A
C 387	13.8	76.7	2518	13	US-10-027-632-103158	Sequence 103158, A
C 388	13.8	76.7	2518	13	US-10-027-632-103157	Sequence 103157, A
C 389	13.8	76.7	2518	13	US-10-027-632-103158	Sequence 103158, A
C 390	13.8	76.7	2568	15	US-10-128-714-488	Sequence 488, App
C 391	13.8	76.7	2683	13	US-10-128-714-488	Sequence 488, App
C 392	13.8	76.7	2718	9	US-09-764-877-3940	Sequence 3940, App
C 393	13.8	76.7	2718	15	US-10-242-515-3940	Sequence 3940, App
C 394	13.8	76.7	3166	16	US-10-108-260A-739	Sequence 739, App
C 395	13.8	76.7	3194	18	US-10-723-860-7737	Sequence 7737, App
C 396	13.8	76.7	3216	17	US-10-437-963-74161	Sequence 74161, A
C 397	13.8	76.7	3858	10	US-09-814-353-21751	Sequence 21751, A
C 398	13.8	76.7	4161	14	US-10-198-846-13097	Sequence 13097, A
C 399	13.8	76.7	4406	17	US-10-437-963-13891	Sequence 13891, A
C 400	13.8	76.7	4518	18	US-10-483-512-73	Sequence 73, App
C 401	13.8	76.7	5825	14	US-10-116-802-123	Sequence 123, App
C 402	13.8	76.7	6393	15	US-10-334-726-323	Sequence 323, App
C 403	13.8	76.7	6394	15	US-10-177-293-363	Sequence 363, App
C 404	13.8	76.7	6415	15	US-10-172-118-1542	Sequence 1542, App
C 405	13.8	76.7	6415	16	US-10-342-887-1542	Sequence 1542, App
C 406	13.8	76.7	6641	18	US-10-723-860-7301	Sequence 7301, App
C 407	13.8	76.7	6930	10	US-09-770-107-1	Sequence 1, App
C 408	13.8	76.7	7450	18	US-10-788-792-11	Sequence 11, App
C 409	13.8	76.7	7450	18	US-10-723-860-1915	Sequence 1915, App
C 410	13.8	76.7	7473	14	US-10-723-860-6240	Sequence 6240, App
C 411	13.8	76.7	8818	14	US-10-175-523-63	Sequence 63, App
C 412	13.8	76.7	9903	9	US-09-764-891-8188	Sequence 8188, App
C 413	13.8	76.7	10785	14	US-09-764-891-824	Sequence 824, App
C 414	13.8	76.7	10785	14	US-10-079-854-324	Sequence 324, App
C 415	13.8	76.7	13315	9	US-09-764-878-328	Sequence 328, App
C 416	13.8	76.7	13315	14	US-10-079-854-328	Sequence 328, App
C 417	13.8	76.7	15654	18	US-10-667-141-58	Sequence 58, App
C 418	13.8	76.7	15654	18	US-10-667-141-59	Sequence 59, App
C 419	13.8	76.7	15654	18	US-10-667-141-60	Sequence 60, App
C 420	13.8	76.7	20261	9	US-09-764-878-325	Sequence 325, App
C 421	13.8	76.7	20261	14	US-10-079-854-325	Sequence 325, App
C 422	13.8	76.7	29328	17	US-10-367-094-189	Sequence 189, App
C 423	13.8	76.7	32835	11	US-09-997-722-214	Sequence 214, App
C 424	13.8	76.7	41079	17	US-10-322-281-516	Sequence 516, App
C 425	13.8	76.7	51001	16	US-10-189-268-11	Sequence 11, App
C 426	13.8	76.7	55996	11	US-09-997-722-7	Sequence 7, App
C 427	13.8	76.7	57013	13	US-10-087-192-1798	Sequence 1798, App
C 428	13.8	76.7	62822	18	US-10-087-192-1753	Sequence 1753, App
C 429	13.8	76.7	67076	13	US-10-087-192-1753	Sequence 1753, App
C 430	13.8	76.7	74580	18	US-10-719-993-6830	Sequence 6830, App
C 431	13.8	76.7	88277	18	US-10-719-993-6931	Sequence 6931, App
C 432	13.8	76.7	89047	13	US-10-672-787-34	Sequence 34, App
C 433	13.8	76.7	95960	16	US-10-087-192-1384	Sequence 1384, App
C 434	13.8	76.7	95960	13	US-10-087-192-1384	Sequence 1384, App
C 435	13.8	76.7	96598	11	US-09-997-722-127	Sequence 127, App
C 436	13.8	76.7	101193	17	US-10-322-281-468	Sequence 468, App
C 437	13.8	76.7	104514	13	US-10-087-192-1361	Sequence 1361, App
C 438	13.8	76.7	122186	10	US-09-563-728A-33	Sequence 36, App
C 439	13.8	76.7	133955	15	US-10-087-192-1244	Sequence 1984, App
C 440	13.8	76.7	136726	13	US-10-085-117-244	Sequence 244, App
C 441	13.8	76.7	154817	17	US-10-085-117-334	Sequence 334, App
C 442	13.8	76.7	168174	15	US-10-071-411-63	Sequence 63, App
C 443	13.8	76.7	168273	17	US-10-071-411-2	Sequence 2, App
C 444	13.8	76.7	200000	17	US-10-672-764A-31	Sequence 31, App
C 445	13.8	76.7	220093	18	US-10-719-993-6861	Sequence 6861, App
C 446	13.8	76.7	230101	18	US-10-719-993-6829	Sequence 6829, App
C 447	13.8	76.7	238910	18	US-10-719-993-6801	Sequence 6801, App
C 448	13.8	76.7	254087	13	US-10-087-192-223	Sequence 223, App
C 449	13.8	76.7	257156	11	US-09-968-007A-427	Sequence 427, App
C 450	13.8	76.7	302603	16	US-10-271-416-8	Sequence 8, App

C 451	13.8	76.7	303172	18	US-10-719-993-6890	Sequence 6890, App
C 452	13.8	76.7	303172	18	US-10-719-993-6890	Sequence 6890, App
C 453	13.8	76.7	374849	13	US-10-087-192-1627	Sequence 1627, App
C 454	13.8	76.7	430442	18	US-10-417-375-128	Sequence 128, App
C 455	13.8	76.7	518360	17	US-10-367-094-125	Sequence 125, App
C 456	13.8	76.7	538165	10	US-09-939-964-1	Sequence 1, App
C 457	13.8	76.7	606398	9	US-10-719-993-6782	Sequence 6782, App
C 458	13.8	76.7	684973	15	US-09-263-959-1	Sequence 1, App
C 459	13.8	76.7	744802	15	US-10-292-798-1369	Sequence 1369, App
C 460	13.8	76.7	775062	13	US-10-719-993-6844	Sequence 6844, App
C 461	13.8	76.7	7940917	18	US-10-027-632-174763	Sequence 174763, App
C 462	13.8	76.7	2940917	15	US-10-027-632-174763	Sequence 174763, App
C 463	13.4	74.4	60	10	US-09-908-975-12684	Sequence 12684, A
C 464	13.4	74.4	190	9	US-10-425-115-104726	Sequence 104726, A
C 465	13.4	74.4	190	9	US-09-560-633-994	Sequence 994, App
C 466	13.4	74.4	201	17	US-10-741-601-25265	Sequence 25265, A
C 467	13.4	74.4	201	18	US-10-719-993-41528	Sequence 41528, A
C 468	13.4	74.4	201	18	US-10-719-993-52468	Sequence 52468, A
C 469	13.4	74.4	216	16	US-10-424-599-123731	Sequence 123731, A
C 470	13.4	74.4	231	17	US-10-469-285-812	Sequence 812, App
C 471	13.4	74.4	293	16	US-10-424-599-121345	Sequence 121345, A
C 472	13.4	74.4	293	16	US-10-425-115-46410	Sequence 46410, A
C 473	13.4	74.4	313	16	US-10-424-599-103027	Sequence 103027, A
C 474	13.4	74.4	326	18	US-10-425-115-21904	Sequence 21904, A
C 475	13.4	74.4	335	18	US-10-425-115-70951	Sequence 70951, A
C 476	13.4	74.4	359	18	US-09-867-701-8048	Sequence 8048, App
C 477	13.4	74.4	359	18	US-10-357-930-55928	Sequence 55928, A
C 478	13.4	74.4	360	11	US-09-732-627A-4038	Sequence 4038, App
C 479	13.4	74.4	361	9	US-09-922-293-2864	Sequence 2864, App
C 480	13.4	74.4	361	11	US-09-922-293-2864	Sequence 2864, App
C 481	13.4	74.4	362	11	US-09-922-293-2864	Sequence 2864, App
C 482	13.4	74.4	365	15	US-10-369-930-28743	Sequence 28743, A
C 483	13.4	74.4	366	11	US-09-922-293-2864	Sequence 2864, App
C 484	13.4	74.4	370	18	US-10-856-499-1783	Sequence 1783, App
C 485	13.4	74.4	370	18	US-10-357-930-21497	Sequence 21497, A
C 486	13.4	74.4	381	18	US-10-357-930-27338	Sequence 27338, A
C 487	13.4	74.4	387	18	US-10-357-930-11517	Sequence 11517, A
C 488	13.4	74.4	395	13	US-10-027-632-138771	Sequence 138771, A
C 489	13.4	74.4	395	15	US-10-027-632-138771	Sequence 138771, A
C 490	13.4	74.4	401	16	US-10-412-698B-1948	Sequence 1948, App
C 491	13.4	74.4	405	16	US-10-424-599-83967	Sequence 83967, A
C 492	13.4	74.4	418	17	US-10-437-963-22447	Sequence 22447, A
C 493	13.4	74.4	422	18	US-10-357-930-32689	Sequence 32689, A
C 494	13.4	74.4	422	11	US-10-357-930-41615	Sequence 41615, A
C 495	13.4	74.4	429	11	US-09-922-293-295	Sequence 295, App
C 496	13.4	74.4	434	18	US-10-425-115-162603	Sequence 162603, A
C 497	13.4	74.4	441	13	US-10-027-632-55822	Sequence 55822, A
C 498	13.4	74.4	441	13	US-10-027-632-296419	Sequence 296419, A
C 499	13.4	74.4	441	15	US-10-027-632-55822	Sequence 55822, A
C 500	13.4	74.4	441	15	US-10-027-632-296419	Sequence 296419, A

ALIGNMENTS

RESULT 1
US-09-917-800A-1491/C
Sequence 1491, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917, 800A
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222, 040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222, 880

PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1491
LENGTH: 1627
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 X13058
US-09-917-800A-1491

Query Match
Best Local Similarity 100.0%; Score 18; DB 9; Length 1627;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18
|||||
Db 41 CGACTGTGATCCTCCAT 24

RESULT 2
US-10-388-934-568/c
Sequence 568, Application US/10388934
Publication No. US20040005547A1
GENERAL INFORMATION:
APPLICANT: Boess, Franziska
APPLICANT: Suter-Dick, Laura
APPLICANT: Wolf, Detlef
TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
FILE REFERENCE: 21199
CURRENT APPLICATION NUMBER: US/10/388,934
CURRENT FILING DATE: 2003-03-14
PRIOR APPLICATION NUMBER: 02005336.9
PRIOR FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 02015657.6
PRIOR FILING DATE: 2002-07-17
NUMBER OF SEQ ID NOS: 862
SOFTWARE: PatentIn version 3.1
SEQ ID NO 568
LENGTH: 1627
TYPE: DNA
ORGANISM: Rattus norvegicus (No. US20040005547A1 way rat)
US-10-388-934-568

Query Match
Best Local Similarity 100.0%; Score 18; DB 16; Length 1627;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18
|||||
Db 41 CGACTGTGATCCTCCAT 24

RESULT 3
US-09-974-300-3673/c
Sequence 3673, Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.

APPLICANT: Clausen, Ib Groth
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
FILE REFERENCE: 10085,500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680,598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3673
LENGTH: 390
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-3673

Query Match
Best Local Similarity 91.1%; Score 16.4; DB 9; Length 390;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18
|||||
Db 120 CGACTGTGATCCTCCAT 103

RESULT 4
US-10-425-115-32582/c
Sequence 32582, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 32582
LENGTH: 724
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1) (724)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: clone ID: MFT4577_129729C.1
US-10-425-115-32582

Query Match
Best Local Similarity 94.4%; Score 16.4; DB 18; Length 724;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18
|||||
Db 176 CGACTGTGATCCTCCAT 159

RESULT 5
US-10-038-010-5/c
Sequence 5, Application US/10038010
Publication No. US20030040089A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
APPLICANT: Pierre, Legrain
TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
FILE REFERENCE: B4767A
CURRENT APPLICATION NUMBER: US/10/038,010

CURRENT FILING DATE: 2002-07-23
PRIOR APPLICATION NUMBER: US 60/259,377
PRIOR FILING DATE: 2001-01-02
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1173
TYPE: DNA
ORGANISM: mouse p53
FEATURE:
NAME/KEY: gene
LOCATION: (1)..(1173)
OTHER INFORMATION: mouse p53 : Tumour suppressor protein
US-10-038-010-5

Query Match 91.1%; Score 16.4; DB 14; Length 1173;
Best Local Similarity 94.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
|||||
Db 27 CGACTGTGAATCCTCCAT 10

RESULT 6
US-10-267-384-214/C
Sequence 214, Application US/10267384
Publication No. US20030198623A1
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
Tataglia, James
Cox, William I.
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/267,384
FILING DATE: 09-Oct-2002
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066CURTMS
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 1173 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-10-267-384-214

Query Match 91.1%; Score 16.4; DB 15; Length 1173;
Best Local Similarity 94.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGAATCCTCCAT 18

Db 27 CGACTGTGAATCCTCCAT 10
|||||

RESULT 7
US-10-746-558-1/C
Sequence 1, Application US/10746558
Publication No. US20040208850A1
GENERAL INFORMATION:
APPLICANT: Ellenhorn, Joshua D.I.
APPLICANT: Diamond, Don J.
TITLE OF INVENTION: Modified vaccinia Ankara expressing p53 in cancer immunotherapy
FILE REFERENCE: 54435,8005,US00
CURRENT APPLICATION NUMBER: US/10/746,558
CURRENT FILING DATE: 2003-12-23
PRIOR APPLICATION NUMBER: 06/436,268
PRIOR FILING DATE: 2002-12-23
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1173
TYPE: DNA
ORGANISM: Mus musculus
US-10-746-558-1

Query Match 91.1%; Score 16.4; DB 18; Length 1173;
Best Local Similarity 94.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
|||||
Db 27 CGACTGTGAATCCTCCAT 10

RESULT 8
US-10-392-113-32/C
Sequence 32, Application US/10392113
Publication No. US20030224993A1
GENERAL INFORMATION:
APPLICANT: Land, Hartmut
APPLICANT: Delieu, Laurent
TITLE OF INVENTION: COMPOSITIONS THAT INHIBIT PROLIFERATION
FILE REFERENCE: 21108,0005U3
CURRENT APPLICATION NUMBER: US/10/392,113
CURRENT FILING DATE: 2003-03-17
PRIOR APPLICATION NUMBER: 60/365,078
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: PCT/US01/32127
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 60/239,705
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 32
LENGTH: 1208
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:/Note =
OTHER INFORMATION: Synthetic Construct
US-10-392-113-32

Query Match 91.1%; Score 16.4; DB 15; Length 1208;
Best Local Similarity 94.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
|||||
Db 52 CGACTGTGAATCCTCCAT 35

RESULT 9

US-10-424-599-102369/c
; Sequence 102369, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 102369
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURES:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63456C.1
US-10-424-599-102369

Query Match 85.6%; Score 15.4; DB 16; Length 441;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18
DB 236 GAATGTGAATCTCCAT 220

RESULT 10
US-10-027-632-285372
; Sequence 285372, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285372
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-285372

Query Match 85.6%; Score 15.4; DB 13; Length 458;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18
DB 235 GACTGTGAATCTCCAT 251

RESULT 11
US-10-027-632-285372
; Sequence 285372, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285372
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-285372

Query Match 85.6%; Score 15.4; DB 15; Length 458;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18
DB 235 GACTGTGAATCTCCAT 251

RESULT 12
US-10-653-047-6563
; Sequence 6563, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653, 047
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533, 559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273, 623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6563
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
US-10-653-047-6563

Query Match 85.6%; Score 15.4; DB 18; Length 584;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GACTGTGAATCCTCCAT 18
|||
Db 240 GATTGTGAATCCTCCAT 256

RESULT 13

US-10-425-115-57974
; Sequence 57974, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 57974
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(604)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: MRT4577_15286C.1
US-10-425-115-57974

Query Match 85.6%; Score 15.4; DB 18; Length 604;
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1;

OY 1 CGACTGTGAATCCTCCA 17
|||
Db 420 CGACTGTGAATCCTCCA 436

RESULT 14

US-10-052-283-349
; Sequence 349, Application US/10052283
; Publication No. US20030064379A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Goddard, Audrey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND METHOD OF USE THEREOF
; FILE REFERENCE: P2751R1C1
; CURRENT APPLICATION NUMBER: US/10/052,283
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US00/20006
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/145,701
; PRIOR FILING DATE: 1999-07-26
; NUMBER OF SEQ ID NOS: 564
; SEQ ID NO 349
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 19, 46, 101, 110, 132, 271
; OTHER INFORMATION: unknown base
US-10-052-283-349

Query Match 85.6%; Score 15.4; DB 14; Length 608;
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1;

OY 1 CGACTGTGAATCCTCCA 17
|||
Db 299 CGACTGTGAATCCTCCA 315

RESULT 15

US-10-425-115-180675
; Sequence 180675, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 180675
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(616)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: MRT4577_96355C.1
US-10-425-115-180675

Query Match 85.6%; Score 15.4; DB 18; Length 616;
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1;

OY 2 GACTGTGAATCCTCCAT 18
|||
Db 262 GACTGTGAATCCTCCAT 278

RESULT 16

US-10-425-115-180677
; Sequence 180677, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 180677
; LENGTH: 853
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 19, 46, 101, 110, 132, 271
; OTHER INFORMATION: unknown base
US-10-425-115-180677

Query Match 85.6%; Score 15.4; DB 18; Length 853;
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1;

OY 2 GACTGTGAATCCTCCAT 18
|||
Db 462 GACTGTGAATCCTCCAT 478

RESULT 17

US-10-425-115-178848
; Sequence 178848, Application US/10425115
; Publication No. US20040214272A1


```

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 178848
; LENGTH: 881
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_94693C.1
US-10-425-115-178848

Query Match      85.6%; Score 15.4; DB 18; Length 881;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      2 GACTGTGAATCCTCCAT 18
      |||||
Db      468 GACTGTGAATCCTCCAT 484

RESULT 18
US-10-425-115-180678
; Sequence 180678, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 180678
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1017)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_96358C.1
US-10-425-115-180678

Query Match      85.6%; Score 15.4; DB 18; Length 1017;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      2 GACTGTGAATCCTCCAT 18
      |||||
Db      475 GACTGTGAATCCTCCAT 491

RESULT 19
US-09-731-872-8
; Sequence 8, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
```

```

; FILE REFERENCE: 78-US3,REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent .pm
; SEQ ID NO 8
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..1398
; NAME/KEY: sig peptide
; LOCATION: 190..252
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.8172934575094
; OTHER INFORMATION: seq ALLWAQEVGVLA/GR
US-09-876-997-8

Query Match      85.6%; Score 15.4; DB 10; Length 1510;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 CGACTGTGAATCCTCCA 17
      |||||
Db      405 CGACTGTGAATCCTCCA 421

RESULT 20
US-09-876-997-8
; Sequence 8, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78-US4,CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent .pm
; SEQ ID NO 8
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..1398
; NAME/KEY: sig peptide
; LOCATION: 190..252
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.8172934575094
; OTHER INFORMATION: seq ALLWAQEVGVLA/GR
US-09-876-997-8

Query Match      85.6%; Score 15.4; DB 10; Length 1510;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 CGACTGTGAATCCTCCA 17
      |||||
Db      405 CGACTGTGAATCCTCCA 421
```

RESULT 21

US-09-731-872-47
; Sequence 47, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78-US3-REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 47
; LENGTH: 1523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 217..1410
; NAME/KEY: sig_peptide
; LOCATION: 217..279
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.8172934575094
; OTHER INFORMATION: seq ALLMAQEVGYLA/GR
US-09-731-872-47

Query Match 85.6%; Score 15.4; DB 9; Length 1523;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCA 17
|||||
Db 432 CGACTGTGAATCCTCCA 448

RESULT 22

US-09-876-997-47
; Sequence 47, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78-US4-CTP
; CURRENT APPLICATION NUMBER: US/09/876,997
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 47
; LENGTH: 1523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 217..1410
; NAME/KEY: sig_peptide
; LOCATION: 217..279
; OTHER INFORMATION: Von Heijne matrix

OTHER INFORMATION: score 5.8172934575094
; OTHER INFORMATION: seq ALLMAQEVGYLA/GR
US-09-876-997-47

Query Match 85.6%; Score 15.4; DB 10; Length 1523;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCA 17
|||||
Db 432 CGACTGTGAATCCTCCA 448

RESULT 23

US-09-935-390A-15
; Sequence 15, Application US/09935390A
; Patent No. US20020076761A1
; GENERAL INFORMATION:
; APPLICANT: Escobedo, Jaime
; Quianjin, Hu
; Garcia, Pablo
; Williams, Lewis T.
; Kothakota, Srinivas
; TITLE OF INVENTION: Secreted Human Proteins
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/935,390A
; FILING DATE: 22-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/988,671
; FILING DATE: 1997-12-11
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. R. Potter
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 1369.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2718
; TELEFAX: (510) 655-3542
; TELEEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1535 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-935-390A-15

Query Match 85.6%; Score 15.4; DB 9; Length 1535;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCA 17
|||||
Db 435 CGACTGTGAATCCTCCA 451

RESULT 24

US-09-745-763-188
; Sequence 188, Application US/09745763

Patent No. US20020065394A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallie, Edward R.
Collins-Racie, Lisa A.
Byans, Cheryl
Merberg, David
Treacy, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 1755 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 188:
US-09-745-763-188
Query Match 85.6%; Score 15.4; DB 9; Length 1755;
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATTCCTCCA 17
Db 673 CGACTGTGATTCCTCCA 689
RESULT 25
US-09-925-298-221
Sequence 221, Application US/09925298
Publication No. US20020039764A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103
CURRENT APPLICATION NUMBER: US/09/925,298
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 221
LENGTH: 1892
TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1892)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-298-221
Query Match 85.6%; Score 15.4; DB 9; Length 1892;
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATTCCTCCA 17
Db 763 CGACTGTGATTCCTCCA 779
RESULT 26
US-10-102-806-221
Sequence 221, Application US/10102806
Publication No. US20030054421A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103P1C1
CURRENT APPLICATION NUMBER: US/10/102,806
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/925,298
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 221
LENGTH: 1892
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1892)
OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-221
Query Match 85.6%; Score 15.4; DB 14; Length 1892;
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATTCCTCCA 17
Db 763 CGACTGTGATTCCTCCA 779
RESULT 27
US-10-322-696-67
Sequence 67, Application US/10322696
Publication No. US20040166490A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
FILE REFERENCE: 529452001200
CURRENT APPLICATION NUMBER: US/10/322,696
CURRENT FILING DATE: 2003-10-17
NUMBER OF SEQ ID NOS: 186
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 67
LENGTH: 154504
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)-(154504)

OTHER INFORMATION: n = A,T,C or G
US-10-322-696-67

Query Match 85.6%; Score 15.4; DB 17; Length 154504;
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
|||||
Db 63891 GACTGTGAATCCTCCAT 63907

RESULT 28
US-10-672-764A-33
; Sequence 33, Application US/10672764A
; Publication No. US20040156832A1
; GENERAL INFORMATION:
; APPLICANT: Jolly, Chris
; TITLE OF INVENTION: Immunoglobulin Compositions and Methods
; FILE REFERENCE: 13311.1001U
; CURRENT APPLICATION NUMBER: US/10/672,764A
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 200000
; TYPE: DNA
; ORGANISM: Human
US-10-672-764A-33

Query Match 85.6%; Score 15.4; DB 17; Length 200000;
Best Local Similarity 94.1%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
|||||
Db 107278 GACTGTGAATCCTCCAT 107294

RESULT 29
US-10-131-786-2
; Sequence 2, Application US/10131786
; Publication No. US20030026831A1
; GENERAL INFORMATION:
; APPLICANT: Lakkaraju, A.
; APPLICANT: Dubinsky, J.M.
; APPLICANT: Low, W.
; TITLE OF INVENTION: Delivery of antisense oligonucleotides to neurons with anionic li
; FILE REFERENCE: 600.519US1
; CURRENT APPLICATION NUMBER: US/10/131,786
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: US 60/285,337
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-131-786-2

Query Match 83.3%; Score 15; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CTGTGAATCCTCCAT 18
|||||
Db 1 CTGTGAATCCTCCAT 15

RESULT 30
US-10-202-193-5/c

Sequence 5, Application US/10202193
; Publication No. US20020192699A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steimann, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; FILE REFERENCE: PP-01532.103/200130.463D1
; CURRENT APPLICATION NUMBER: US/10/202,193
; CURRENT FILING DATE: 2002-07-23
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc. Feature
; LOCATION: (1) .. (1024)
; OTHER INFORMATION: n = A,T,C or G
US-10-202-193-5

Query Match 83.3%; Score 15; DB 13; Length 1024;
Best Local Similarity 93.8%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACTGTGAATCCTCCAT 18
|||||
Db 574 ACTGTGAATCCTCCAT 559

RESULT 31
US-10-417-375-128
; Sequence 128, Application US/10417375
; Publication No. US20040219528A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001600
; CURRENT APPLICATION NUMBER: US/10/417,375
; CURRENT FILING DATE: 2003-04-15
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128
; LENGTH: 430442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-417-375-128

Query Match 83.3%; Score 15; DB 18; Length 430442;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ACTGTGAATCCTCCA 17
|||||
Db 139298 ACTGTGAATCCTCCA 139312

RESULT 32
US-10-425-115-65084/c
; Sequence 65084, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53222)B
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 65084
LENGTH: 225
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MNT4577_159354C.1
US-10-425-115-65084

Query Match 82.2% Score 14.8; DB 18; Length 225;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCCAT 18
DB 88 CGTCTGTGATCTCCAT 71

RESULT 33
US-09-922-293-1999
Sequence 1999, Application US/09922293
Publication No. US2004012339A1
GENERAL INFORMATION:
APPLICANT: Comer, Timothy W.
APPLICANT: Heck, Gregory R.
APPLICANT: Liu, Jingdong
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 16517.254
CURRENT FILING DATE: US/09/922,293
CURRENT FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: US 60/067,000
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: US 60/069,472
PRIOR FILING DATE: 1997-12-09
PRIOR APPLICATION NUMBER: US 60/071,479
PRIOR FILING DATE: 1998-01-13
PRIOR APPLICATION NUMBER: US 60/074,201
PRIOR FILING DATE: 1998-02-10
PRIOR APPLICATION NUMBER: US 60/074,282
PRIOR FILING DATE: 1998-02-10
PRIOR APPLICATION NUMBER: US 60/074,280
PRIOR FILING DATE: 1998-02-10
PRIOR APPLICATION NUMBER: US 60/074,281
PRIOR FILING DATE: 1998-02-10
PRIOR APPLICATION NUMBER: US 60/074,566
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: US 60/074,567
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: US 60/074,565
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: US 60/075,462
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: US 60/074,789
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: US 60/075,459
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: US 60/075,461
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: US 60/075,464
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: US 60/075,460
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: US 60/075,463
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: US 60/077,231
PRIOR FILING DATE: 1998-03-09

PRIOR APPLICATION NUMBER: US 60/077,229
PRIOR FILING DATE: 1998-03-09
PRIOR APPLICATION NUMBER: US 60/077,230
PRIOR FILING DATE: 1998-03-09
PRIOR APPLICATION NUMBER: US 60/078,368
PRIOR FILING DATE: 1998-03-18
PRIOR APPLICATION NUMBER: US 60/080,844
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 60/083,067
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: US 60/083,386
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: US 60/083,387
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: US 60/083,388
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: US 60/083,389
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: US 60/085,224
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: US 60/085,223
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: US 60/085,222
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: US 60/085,533
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: US 60/086,186
PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/086,187
PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/086,185
PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/086,184
PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/086,183
PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/086,188
PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/089,524
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: US 60/089,810
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,814
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,808
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,812
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,807
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,806
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,813
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,811
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,793
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/091,405
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/091,247
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/099,667
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/099,668
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/099,670
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/099,697
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/100,674
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: US 60/100,673

;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: US 60/100,672
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: US 60/100,963
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: US 60/101,131
;; PRIOR FILING DATE: 1998-09-21
;; PRIOR APPLICATION NUMBER: US 60/101,132
;; PRIOR FILING DATE: 1998-09-21
;; PRIOR APPLICATION NUMBER: US 60/101,130
;; PRIOR FILING DATE: 1998-09-21
;; PRIOR APPLICATION NUMBER: US 60/101,508
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: US 60/101,344
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: US 60/101,347
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: US 60/101,343
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: US 60/101,707
;; PRIOR FILING DATE: 1998-09-25
;; PRIOR APPLICATION NUMBER: US 60/104,126
;; PRIOR FILING DATE: 1998-10-13
;; PRIOR APPLICATION NUMBER: US 60/104,128
;; PRIOR FILING DATE: 1998-10-13
;; PRIOR APPLICATION NUMBER: US 60/104,127
;; PRIOR FILING DATE: 1998-10-13
;; PRIOR APPLICATION NUMBER: US 60/104,124
;; PRIOR FILING DATE: 1998-10-13
;; PRIOR APPLICATION NUMBER: US 60/109,018
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: US 60/108,996
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: US 09/199,129
;; PRIOR FILING DATE: 1998-11-24
;; PRIOR APPLICATION NUMBER: US 09/210,297
;; PRIOR FILING DATE: 1998-12-08
;; PRIOR APPLICATION NUMBER: US 60/111,981
;; PRIOR FILING DATE: 1998-12-11
;; PRIOR APPLICATION NUMBER: US 60/113,224
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: US 09/229,413
;; PRIOR FILING DATE: 1999-01-12
;; NUMBER OF SEQ ID NOS: 3853
;; SEQ ID NO 1999
;; LENGTH: 281
;; TYPE: DNA
;; ORGANISM: Glycine max
US-09-922-293-1999

Query Match 82.2%; Score 14.8; DB 11; Length 281;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGACTGGAATCCTCCAT 18
Db 258 CGAGTGTGATCCTCCAT 275
RESULT 34
US-10-242-535A-30560/C
; Sequence 30560, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340

;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: US 60/275,017
;; PRIOR FILING DATE: 2001-03-12
;; PRIOR APPLICATION NUMBER: US 60/271,955
;; PRIOR FILING DATE: 2001-02-28
;; NUMBER OF SEQ ID NOS: 58994
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 30560
;; LENGTH: 350
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (340)..(340)
;; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-30560

Query Match 82.2%; Score 14.8; DB 16; Length 350;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGACTGGAATCCTCCAT 18
Db 72 CGACTGGAATCCACAT 55
RESULT 35
US-10-085-783A-30560/C
; Sequence 30560, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30560
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (340)..(340)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-30560

Query Match 82.2%; Score 14.8; DB 16; Length 350;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGACTGGAATCCTCCAT 18
Db 72 CGACTGGAATCCACAT 55
RESULT 36
US-09-954-456-1075
; Sequence 1075, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cance
; FILE REFERENCE: 689290-76

```
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1075
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1075
```

```
Query Match      82.2%; Score 14.8; DB 9; Length 417;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CGACTGTGAATCCTCCAT 18
Db      278 CGACTGTGAATCCTCCAT 295
```

```
RESULT 37
US-09-954-456-1774
; Sequence 1774, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
```

```
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 1774
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1774
```

```
Query Match      82.2%; Score 14.8; DB 9; Length 417;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CGACTGTGAATCCTCCAT 18
Db      278 CGACTGTGAATCCTCCAT 295
```

```
RESULT 38
US-09-880-107-81
; Sequence 81, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Logic, Inc.
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA035540
; NAME/KEY: unsure
; LOCATION: (1)...(417)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-81
```

```
Query Match      82.2%; Score 14.8; DB 9; Length 417;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 CGACTGTGAATCCTCCAT 18
Db      278 CGACTGTGAATCCTCCAT 295
```

```
RESULT 39
US-09-918-995-19674/c
; Sequence 19674, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
```

```
SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19674
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1) ..(474)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-19674

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 10; Length 474;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
Db 343 CGACTGTGAATCTCCAT 326

RESULT 40
US-10-027-632-137650
; Sequence 137650, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 137650
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-137650

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 13; Length 578;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
Db 423 CGACTGTGAATCTCCAT 440

RESULT 41
US-10-027-632-137650
; Sequence 137650, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
```

```
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 137650
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-137650

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 15; Length 578;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
Db 423 CGACTGTGAATCTCCAT 440

RESULT 42
US-10-425-115-47892
; Sequence 47892, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 47892
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_143684C.1
US-10-425-115-47892

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 18; Length 600;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
Db 396 CGACTGTGAATCTCCAT 413

RESULT 43
US-10-767-701-519/C
; Sequence 519, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2002-04-30
```


RESULT 45
 US-10-260-238-3388
 : Sequence 3388, Application US/10260238
 : Publication No. US20040016025A1
 GENERAL INFORMATION:
 :
 : APPLICANT: Budworth, Paul R.
 : APPLICANT: Moughamer, Todd G.
 : APPLICANT: Brigg8, Steven P.
 : APPLICANT: Cooper, Bret
 : APPLICANT: Glazebrook, Jane
 : APPLICANT: Goff, Stephen A.
 : APPLICANT: Katagiri, Fumiyaki
 : APPLICANT: Kreps, Joel
 : APPLICANT: Procart, Nicholas
 : APPLICANT: Ricke, Darrell
 : APPLICANT: Zhu, Jiong
 : TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
 : FILE REFERENCE: 60111-NP
 : CURRENT APPLICATION NUMBER: US/10/260,238
 : CURRENT FILING DATE: 2002-09-26

```

, RESULT 47
, US-10-437-963-34632
, Sequence 34632, Application US/10437963
, Publication No. US2004012343A1
, GENERAL INFORMATION:
, APPLICANT: La Rosa, Thomas J.
, APPLICANT: Kovalic, David K.
, APPLICANT: Zhou, Yihua
, APPLICANT: Cao, Yongwei
, APPLICANT: Wu, Wei
, APPLICANT: Boukharov, Andrey A.
, APPLICANT: Barbazuk, Brad
, APPLICANT: Li, Ping
, TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
, TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
, FILE REFERENCE: 38-21(53221)B
, CURRENT APPLICATION NUMBER: US/10/437, 963
, CURRENT FILING DATE: 2003-05-14
, NUMBER OF SEQ ID NOS: 204966
, SEQ ID NO 34632
, LENGTH: 977

```

TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_38629C.1
US-10-437-963-34632

Query Match 82.2%; Score 14.8; DB 16; Length 977;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
| | | | | | | | | | | | | | | | | | | | | |
Db 636 CCACTGTGTATCCTCCAT 653

RESULT 48
US-10-258-662-58/c

Sequence 58, Application US/10258662
Publication No. US20040082029A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: LAL, Preeti

APPLICANT: YUE, Henry

APPLICANT: TANG, Y. Tom

APPLICANT: LU, Dyrung Aina M.

APPLICANT: AZIMZAI, Yalda

APPLICANT: AU-YOUNG, Janice

APPLICANT: HITLMAN, Jennifer L.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: YAO, Monique G.

APPLICANT: BURFORD, Neil

APPLICANT: BATRA, Sateev

APPLICANT: POLICKY, Jennifer J.

TITLE OF INVENTION: RNA METABOLISM PROTEINS

FILE REFERENCE: PR-0771 USN

CURRENT APPLICATION NUMBER: US/10/258, 662

PRIOR FILING DATE: 2002-10-25

PRIOR APPLICATION NUMBER: US 60/201, 875

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/200, 164

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/202, 090

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/210, 232

PRIOR FILING DATE: 2000-06-06

PRIOR APPLICATION NUMBER: US 60/220, 553

PRIOR FILING DATE: 2000-07-25

NUMBER OF SEQ ID NOS: 94

SOFTWARE: PERL Program

SEQ ID NO 58

LENGTH: 1188

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No: 2099829CB1
US-10-258-662-58

Query Match 82.2%; Score 14.8; DB 16; Length 1188;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
| | | | | | | | | | | | | | | | | | | | | |
Db 666 CGACTGTGAATCCTCCAT 649

RESULT 49

US-10-424-599-82180/c

Sequence 82180, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424, 599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 82180

LENGTH: 1448

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_45225C.1
US-10-424-599-82180

Query Match 82.2%; Score 14.8; DB 16; Length 1448;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
| | | | | | | | | | | | | | | | | | | | | |
Db 628 CCACTGTGAATCCTCCAT 611

RESULT 50
US-10-437-963-35984

Sequence 35984, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/437, 963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 35984

LENGTH: 1826

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_39853C.1
US-10-437-963-35984

Query Match 82.2%; Score 14.8; DB 17; Length 1826;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
| | | | | | | | | | | | | | | | | | | | | |
Db 1360 CGACTGTGAATCCTCCAT 1377

Search completed: January 14, 2005, 19:47:52
Job time : 358.579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 14:47:55 ; Search time 63.9474 Seconds
(without alignments)
200.074 Million cell updates/sec

Title: US-09-578-453-1

Sequence: 1 CGACTGTCATCCTCCAT 18

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTCTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfilest.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18	100.0	1415	3	US-09-414-436-2
C 2	16.4	91.1	1173	2	US-08-184-009-214
C 3	16.4	91.1	1173	2	US-08-458-356-214
C 4	16.4	91.1	1173	3	US-08-796-101-45
C 5	16.4	91.1	1173	3	US-08-460-736-214
C 6	16.4	91.1	1173	3	US-09-535-370-214
C 7	16.4	91.1	1173	4	US-09-663-667-214
C 8	15.4	85.6	396	4	US-08-585-593A-36
C 9	15.4	85.6	527	4	US-09-621-976-1600
C 10	15.4	85.6	2516	4	US-09-270-767-1686
C 11	15	83.3	1024	4	US-09-328-475C-5
C 12	14.8	82.2	1107	4	US-09-248-796A-8906
C 13	14.4	80.0	478	4	US-09-621-976-18989
C 14	14.4	80.0	1093	3	US-09-186-276B-53
C 15	14.4	80.0	1093	4	US-08-847-445-53
C 16	14.4	80.0	1093	4	US-09-186-188B-53
C 17	14.4	80.0	1629	4	US-09-919-039-32
C 18	14.4	80.0	1850	3	US-08-724-468B-3
C 19	14.4	80.0	1850	3	US-08-882-164D-3
C 20	14.4	80.0	3152	4	US-09-399-081A-7
C 21	14.4	80.0	6911	1	US-08-311-174-4
C 22	14.4	80.0	8878	1	US-08-205-176-3
C 23	14	77.8	1395	4	US-09-614-221A-167
C 24	13.8	76.7	432	4	US-09-513-999C-36262
C 25	13.8	76.7	514	4	US-09-389-681-417
C 26	13.8	76.7	514	4	US-09-620-405B-417
C 27	13.8	76.7	514	4	US-09-433-826B-417

C 28	13.8	76.7	514	4	US-09-604-287A-417	Sequence 417, App
C 29	13.8	76.7	514	4	US-09-834-759-417	Sequence 417, App
C 30	13.8	76.7	514	4	US-09-530-751A-417	Sequence 417, App
C 31	13.8	76.7	516	4	US-09-489-039A-2781	Sequence 2781, App
C 32	13.8	76.7	519	4	US-09-389-681-258	Sequence 258, App
C 33	13.8	76.7	519	4	US-09-620-405B-258	Sequence 258, App
C 34	13.8	76.7	519	4	US-09-339-338-258	Sequence 258, App
C 35	13.8	76.7	519	4	US-09-433-826B-258	Sequence 258, App
C 36	13.8	76.7	519	4	US-09-604-287A-258	Sequence 258, App
C 37	13.8	76.7	519	4	US-09-834-759-258	Sequence 258, App
C 38	13.8	76.7	519	4	US-09-590-751A-258	Sequence 258, App
C 39	13.8	76.7	598	4	US-09-270-767-11839	Sequence 11839, A
C 40	13.8	76.7	1041	4	US-09-371-307-79	Sequence 79, App
C 41	13.8	76.7	1146	3	US-09-257-580-1	Sequence 1, App
C 42	13.8	76.7	1146	3	US-09-257-580-4	Sequence 4, App
C 43	13.8	76.7	1166	4	US-09-673-335A-57	Sequence 57, App
C 44	13.8	76.7	1270	4	US-09-270-767-12871	Sequence 12871, A
C 45	13.8	76.7	1431	4	US-10-000-489-87	Sequence 87, App
C 46	13.8	76.7	1431	4	US-10-000-489-89	Sequence 89, App
C 47	13.8	76.7	1458	4	US-09-248-786A-5727	Sequence 5727, App
C 48	13.8	76.7	1518	4	US-09-719-919A-2	Sequence 2, App
C 49	13.8	76.7	1734	3	US-08-687-590-60	Sequence 60, App
C 50	13.8	76.7	1938	4	US-09-673-335A-140	Sequence 140, App
C 51	13.8	76.7	1955	4	US-09-270-767-13672	Sequence 13672, A
C 52	13.8	76.7	2070	4	US-09-540-236-1766	Sequence 1766, App
C 53	13.8	76.7	2307	3	US-09-774-528-314	Sequence 314, App
C 54	13.8	76.7	4377	3	US-08-911-853-28	Sequence 28, App
C 55	13.8	76.7	4377	3	US-09-479-409-28	Sequence 28, App
C 56	13.8	76.7	4377	3	US-09-479-453-28	Sequence 28, App
C 57	13.8	76.7	5240	4	US-09-171-337A-2	Sequence 2, App
C 58	13.8	76.7	5240	4	US-09-631-022-2	Sequence 2, App
C 59	13.8	76.7	89047	4	US-09-556-002-34	Sequence 34, App
C 60	13.8	76.7	168174	4	US-10-071-411A-63	Sequence 63, App
C 61	13.8	76.7	168273	4	US-10-071-411A-2	Sequence 2, App
C 62	13.8	76.7	536165	4	US-09-214-808-1	Sequence 1, App
C 63	13.4	74.4	79	3	US-08-921-887-42	Sequence 42, App
C 64	13.4	74.4	392	4	US-09-270-767-77497	Sequence 27497, A
C 65	13.4	74.4	478	4	US-09-270-767-8209	Sequence 8209, App
C 66	13.4	74.4	478	4	US-09-270-767-33491	Sequence 23491, A
C 67	13.4	74.4	600	4	US-09-621-976-10493	Sequence 10493, A
C 68	13.4	74.4	611	4	US-09-270-767-1675	Sequence 2675, App
C 69	13.4	74.4	611	4	US-09-270-767-19957	Sequence 19957, A
C 70	13.4	74.4	769	4	US-09-513-775B-5	Sequence 5, App
C 71	13.4	74.4	907	4	US-09-513-775B-3	Sequence 3, App
C 72	13.4	74.4	917	4	US-09-270-767-11849	Sequence 11849, A
C 73	13.4	74.4	918	1	US-08-328-322-3	Sequence 3, App
C 74	13.4	74.4	956	4	US-09-641-638-36	Sequence 36, App
C 75	13.4	74.4	956	4	US-09-641-638-37	Sequence 37, App
C 76	13.4	74.4	956	4	US-09-641-638-38	Sequence 38, App
C 77	13.4	74.4	956	4	US-09-641-638-39	Sequence 39, App
C 78	13.4	74.4	956	4	US-09-641-638-40	Sequence 40, App
C 79	13.4	74.4	956	4	US-10-170-097-36	Sequence 36, App
C 80	13.4	74.4	956	4	US-10-170-097-37	Sequence 37, App
C 81	13.4	74.4	956	4	US-10-170-097-38	Sequence 38, App
C 82	13.4	74.4	956	4	US-10-170-097-39	Sequence 39, App
C 83	13.4	74.4	956	4	US-10-170-097-40	Sequence 40, App
C 84	13.4	74.4	1001	4	US-09-641-638-428	Sequence 628, App
C 85	13.4	74.4	1102	4	US-10-170-097-428	Sequence 628, App
C 86	13.4	74.4	1158	4	US-09-107-532A-2119	Sequence 2119, App
C 87	13.4	74.4	1158	4	US-09-328-352-024	Sequence 2024, App
C 88	13.4	74.4	1191	3	US-09-134-001C-1709	Sequence 1709, App
C 89	13.4	74.4	1206	4	US-09-248-786A-3	Sequence 3, App
C 90	13.4	74.4	1222	1	US-08-328-322-1	Sequence 1, App
C 91	13.4	74.4	1281	4	US-09-620-312D-849	Sequence 849, App
C 92	13.4	74.4	1418	4	US-09-673-335A-26	Sequence 26, App
C 93	13.4	74.4	1430	2	US-08-743-130A-1	Sequence 1, App
C 94	13.4	74.4	2024	4	US-09-398-395A-25	Sequence 25, App
C 95	13.4	74.4	2024	4	US-09-887-586A-25	Sequence 25, App
C 96	13.4	74.4	2024	4	US-09-895-752-25	Sequence 25, App
C 97	13.4	74.4	2024	4	US-09-903-012B-25	Sequence 25, App
C 98	13.4	74.4	2024	4	US-09-900-797-25	Sequence 25, App
C 99	13.4	74.4	2025	2	US-08-937-540-1	Sequence 1, App
C 100	13.4	74.4	2745	1	US-08-328-322-6	Sequence 6, App

101	13.4	74.4	2748	4	US-09-614-221A-288	Sequence 288, App	174	12.8	71.1	36	1	US-08-403-762A-153	Sequence 153, App
102	13.4	74.4	3288	1	US-08-208-008C-3	Sequence 3, Appli	175	12.8	71.1	61	1	US-07-744-282C-111	Sequence 111, App
103	13.4	74.4	3500	2	US-08-328-322-4	Sequence 4, Appli	176	12.8	71.1	61	1	US-07-744-282C-113	Sequence 113, App
104	13.4	74.4	11444	2	US-08-222-617A-26	Sequence 26, Appli	177	12.8	71.1	61	5	PCT-US92-06821A-57	Sequence 57, Appli
105	13.4	74.4	12364	2	US-08-222-617A-1	Sequence 1, Appli	178	12.8	71.1	61	5	PCT-US92-06821A-59	Sequence 59, Appli
106	13.4	74.4	20674	4	US-09-641-638-651	Sequence 651, App	179	12.8	71.1	236	4	US-09-513-999C-12376	Sequence 12376, A
107	13.4	74.4	20674	4	US-10-170-097-651	Sequence 651, App	180	12.8	71.1	312	3	US-09-188-930-21	Sequence 21, Appli
108	13.4	74.4	41708	3	US-09-470-512A-3	Sequence 3, Appli	181	12.8	71.1	312	4	US-09-312-283C-21	Sequence 21, Appli
109	13.4	74.4	41708	4	US-09-676-539-18	Sequence 18, Appli	182	12.8	71.1	315	3	US-09-060-756-127	Sequence 127, App
110	13.4	74.4	99916	4	US-09-816-095-3	Sequence 3, Appli	183	12.8	71.1	315	4	US-09-670-314-127	Sequence 127, App
111	13.4	74.4	786431	4	US-09-751-389-3	Sequence 3, Appli	184	12.8	71.1	324	4	US-09-328-352-995	Sequence 995, App
112	13.2	73.3	296	4	US-09-107-532A-525	Sequence 525, App	185	12.8	71.1	335	3	US-08-905-223-194	Sequence 194, App
113	13.2	73.3	246	4	US-09-313-294A-971	Sequence 971, App	186	12.8	71.1	351	4	US-09-513-999C-30457	Sequence 30457, A
114	13.2	73.3	290	4	US-09-612-402B-11	Sequence 11, Appli	187	12.8	71.1	360	4	US-09-543-681A-6	Sequence 6, Appli
115	13.2	73.3	474	1	US-07-845-592J-5	Sequence 5, Appli	188	12.8	71.1	370	4	US-09-513-999C-24920	Sequence 24920, A
116	13.2	73.3	657	4	US-09-084-303B-213	Sequence 213, App	189	12.8	71.1	384	4	US-09-513-999C-27193	Sequence 27193, A
117	13.2	73.3	780	4	US-09-246-796A-3102	Sequence 3102, Ap	190	12.8	71.1	398	4	US-09-621-976-10095	Sequence 10095, A
118	13.2	73.3	831	4	US-09-904-615-25	Sequence 25, Appli	191	12.8	71.1	409	4	US-09-270-767-2400	Sequence 2400, Ap
119	13.2	73.3	867	4	US-09-252-991A-6957	Sequence 6957, Ap	192	12.8	71.1	409	4	US-09-270-767-17682	Sequence 17682, A
120	13.2	73.3	888	4	US-09-134-000C-2240	Sequence 2240, Ap	193	12.8	71.1	423	4	US-09-248-796A-2788	Sequence 2788, Ap
121	13.2	73.3	1166	1	US-07-845-592J-3	Sequence 3, Appli	194	12.8	71.1	424	4	US-09-621-976-17245	Sequence 17245, A
122	13.2	73.3	1444	4	US-09-802-927-2	Sequence 2, Appli	195	12.8	71.1	446	4	US-09-621-976-18000	Sequence 18000, A
123	13.2	73.3	1444	4	US-09-612-402B-11	Sequence 11, Appli	196	12.8	71.1	510	4	US-09-621-976-17134	Sequence 17134, A
124	13.2	73.3	1614	4	US-09-134-000C-926	Sequence 926, App	197	12.8	71.1	513	4	US-09-248-796A-6137	Sequence 6137, Ap
125	13.2	73.3	1656	4	US-09-248-796A-608	Sequence 608, App	198	12.8	71.1	534	4	US-09-270-767-5250	Sequence 5250, Ap
126	13.2	73.3	1734	4	US-09-248-796A-4005	Sequence 4005, Ap	199	12.8	71.1	535	4	US-09-270-767-20532	Sequence 20532, A
127	13.2	73.3	2038	3	US-08-868-423-6	Sequence 45, Appli	200	12.8	71.1	622	4	US-09-621-976-17866	Sequence 17866, Ap
128	13.2	73.3	2052	4	US-08-635-552A-1	Sequence 6, Appli	201	12.8	71.1	632	4	US-09-270-767-3402	Sequence 3402, Ap
129	13.2	73.3	2116	4	US-09-155-768-1	Sequence 1, Appli	202	12.8	71.1	632	4	US-09-270-767-18684	Sequence 18684, A
130	13.2	73.3	2235	3	US-09-248-796A-3987	Sequence 3987, Ap	203	12.8	71.1	655	3	US-09-328-111-727	Sequence 2789, Ap
131	13.2	73.3	2520	4	US-09-799-451-4	Sequence 4, Appli	204	12.8	71.1	667	4	US-09-270-767-7197	Sequence 7197, Ap
132	13.2	73.3	2601	4	US-09-799-451-3	Sequence 3, Appli	205	12.8	71.1	711	6	US-09-270-767-22479	Sequence 22479, A
133	13.2	73.3	2728	4	US-09-620-312D-572	Sequence 572, App	206	12.8	71.1	711	6	5169941-24	Patent No. 5169941
134	13.2	73.3	2949	4	US-09-556-877-170	Sequence 170, App	207	12.8	71.1	711	6	US-08-470-535-13	Sequence 13, Appli
135	13.2	73.3	2949	4	US-09-620-412C-170	Sequence 170, App	208	12.8	71.1	729	3	US-09-270-767-77537	Sequence 77537, Ap
136	13.2	73.3	2949	4	US-09-598-419-170	Sequence 170, App	209	12.8	71.1	736	4	US-09-270-767-23039	Sequence 23039, A
137	13.2	73.3	3021	4	US-09-556-877-182	Sequence 182, App	210	12.8	71.1	737	4	US-09-270-767-14611	Sequence 14611, A
138	13.2	73.3	3021	4	US-09-620-412C-182	Sequence 182, App	211	12.8	71.1	765	4	US-09-673-395A-55	Sequence 55, Appli
139	13.2	73.3	3021	4	US-09-598-419-182	Sequence 7, Appli	212	12.8	71.1	774	4	US-09-107-532A-2311	Sequence 2311, Ap
140	13.2	73.3	3324	4	US-09-612-402B-24	Sequence 24, Appli	213	12.8	71.1	786	2	US-08-644-664B-29	Sequence 29, Appli
141	13.2	73.3	3354	4	US-09-612-402B-21	Sequence 21, Appli	214	12.8	71.1	789	4	US-09-248-796A-2661	Sequence 2661, Ap
142	13.2	73.3	3529	4	US-09-799-451-721	Sequence 721, App	215	12.8	71.1	798	6	5169941-23	Patent No. 5169941
143	13.2	73.3	3738	1	US-08-158-232-7	Sequence 7, Appli	216	12.8	71.1	813	4	US-09-107-532A-3565	Sequence 3565, Ap
144	13.2	73.3	3738	1	US-08-304-626-7	Sequence 7, Appli	217	12.8	71.1	823	3	US-09-799-451-197	Sequence 197, App
145	13.2	73.3	3738	1	US-08-611-928-7	Sequence 7, Appli	218	12.8	71.1	861	4	US-08-998-416-505	Sequence 505, App
146	13.2	73.3	3738	3	US-09-173-891-7	Sequence 7, Appli	219	12.8	71.1	863	4	US-09-756-983-23	Sequence 23, Appli
147	13.2	73.3	4354	3	US-09-802-927-1	Sequence 1, Appli	220	12.8	71.1	870	4	US-09-673-395A-540	Sequence 540, App
148	13.2	73.3	4415	3	US-09-486-580A-1	Sequence 1, Appli	221	12.8	71.1	913	2	US-08-975-316-61	Sequence 61, Appli
149	13.2	73.3	4435	4	US-09-612-402B-1	Sequence 1, Appli	222	12.8	71.1	913	4	US-09-615-192A-61	Sequence 61, Appli
150	13.2	73.3	7431	3	US-09-306-998-2	Sequence 2, Appli	223	12.8	71.1	933	4	US-09-169-789-61	Sequence 61, Appli
151	13.2	73.3	7650	4	US-09-221-017B-911	Sequence 911, App	224	12.8	71.1	997	1	US-08-480-784-25	Sequence 25, Appli
152	13.2	73.3	45716	3	US-08-965-048-5	Sequence 5, Appli	225	12.8	71.1	997	1	US-08-483-553-25	Sequence 25, Appli
153	13.2	73.3	45889	3	US-09-103-840A-2	Sequence 2, Appli	226	12.8	71.1	997	1	US-08-487-002-25	Sequence 25, Appli
154	13.2	73.3	4403765	3	US-09-103-840A-1	Sequence 1, Appli	227	12.8	71.1	997	1	US-08-483-554B-25	Sequence 25, Appli
155	13.2	73.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli	228	12.8	71.1	997	1	US-08-488-011B-25	Sequence 25, Appli
156	13.2	73.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli	229	12.8	71.1	997	1	US-08-850-727-25	Sequence 25, Appli
157	13.2	73.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli	230	12.8	71.1	997	1	PCT-US95-10203-25	Sequence 25, Appli
158	13.2	73.3	544	4	US-09-673-395A-48	Sequence 48, Appli	231	12.8	71.1	997	1	PCT-US95-10203-25	Sequence 25, Appli
159	13.2	73.3	869	4	US-10-140-002-41	Sequence 41, Appli	232	12.8	71.1	1089	4	US-09-248-796A-7664	Sequence 7664, Ap
160	13.2	73.3	1027	4	US-09-684-405-8	Sequence 8, Appli	233	12.8	71.1	1116	1	US-08-244-646-14	Sequence 14, Appli
161	13.2	73.3	1825	4	US-09-270-767-14929	Sequence 14929, A	234	12.8	71.1	1141	3	US-09-147-550-10	Sequence 10, Appli
162	13.2	73.3	3001	4	US-09-539-333D-180	Sequence 180, App	235	12.8	71.1	1141	3	US-09-557-917-10	Sequence 10, Appli
163	13.2	73.3	3001	4	US-08-911-853-5	Sequence 5, Appli	236	12.8	71.1	1144	3	US-09-257-179-15	Sequence 15, Appli
164	12.8	71.1	19	3	US-09-479-409-5	Sequence 5, Appli	237	12.8	71.1	1160	6	5169941-2	Patent No. 5169941
165	12.8	71.1	19	3	US-09-479-409-5	Sequence 5, Appli	238	12.8	71.1	1192	4	US-09-566-921-122	Sequence 122, App
166	12.8	71.1	19	3	US-09-479-409-5	Sequence 5, Appli	239	12.8	71.1	1200	4	US-09-614-221A-31	Sequence 31, Appli
167	12.8	71.1	21	4	US-08-649-950-56	Sequence 56, Appli	240	12.8	71.1				
168	12.8	71.1	23	4	US-09-242-632A-3	Sequence 3, Appli	241	12.8	71.1				
169	12.8	71.1	34	5	PCT-US95-07554-1	Sequence 1, Appli	242	12.8	71.1				
170	12.8	71.1	35	1	US-07-744-282C-110	Sequence 110, App	243	12.8	71.1				
171	12.8	71.1	35	1	US-07-744-282C-112	Sequence 112, App	244	12.8	71.1				
172	12.8	71.1	35	1	PCT-US92-06821A-56	Sequence 56, Appli	245	12.8	71.1				
173	12.8	71.1	35	5	PCT-US92-06821A-58	Sequence 58, Appli	246	12.8	71.1				

C 247	12.8	71.1	1233	4	US-09-543-681A-99	Sequence 99, Appl	330	12.8	71.1	4594	3	US-08-796-101-9	Sequence 9, Appl
C 248	12.8	71.1	1249	4	US-09-495-050A-187	Sequence 187, App	321	12.8	71.1	4594	3	US-09-085-273-43	Sequence 43, Appl
C 249	12.8	71.1	1283	4	US-09-270-767-11737	Sequence 11737, A	322	12.8	71.1	4594	3	US-09-085-273-45	Sequence 45, Appl
C 250	12.8	71.1	1307	2	US-08-078-311-17	Sequence 17, Appl	323	12.8	71.1	4594	4	US-09-916-963-43	Sequence 43, Appl
C 251	12.8	71.1	1307	2	US-08-460-402-17	Sequence 17, Appl	324	12.8	71.1	4594	4	US-09-916-963-45	Sequence 45, Appl
C 252	12.8	71.1	1344	4	US-09-328-352-3791	Sequence 3791, Ap	325	12.8	71.1	4768	2	US-08-658-665-41	Sequence 41, Appl
C 253	12.8	71.1	1350	4	US-09-799-451-799	Sequence 799, App	326	12.8	71.1	4768	3	US-08-796-101-5	Sequence 5, Appl
C 254	12.8	71.1	1365	4	US-09-023-655-119	Sequence 119, App	327	12.8	71.1	4768	3	US-09-085-273-41	Sequence 41, Appl
C 255	12.8	71.1	1412	4	US-09-615-192A-129	Sequence 129, App	328	12.8	71.1	4768	3	US-09-916-963-41	Sequence 41, Appl
C 256	12.8	71.1	1412	4	US-09-169-789-129	Sequence 129, App	329	12.8	71.1	5062	2	US-08-658-665-187	Sequence 187, App
C 257	12.8	71.1	1448	4	US-09-599-360B-3	Sequence 3, Appl	330	12.8	71.1	5183	2	US-08-796-101-42	Sequence 42, Appl
C 258	12.8	71.1	1448	4	US-09-599-360B-49	Sequence 49, Appl	331	12.8	71.1	5183	2	US-08-459-568-3	Sequence 3, Appl
C 259	12.8	71.1	1523	1	US-08-184-327A-7	Sequence 7, Appl	332	12.8	71.1	5183	2	US-08-399-411-3	Sequence 3, Appl
C 260	12.8	71.1	1523	5	PCT-US95-00670-7	Sequence 7, Appl	333	12.8	71.1	5275	3	US-08-796-101-49	Sequence 49, Appl
C 261	12.8	71.1	1619	4	US-09-620-312D-749	Sequence 749, App	334	12.8	71.1	5359	2	US-09-070-060-1	Sequence 1, Appl
C 262	12.8	71.1	1624	4	US-10-140-002-181	Sequence 181, App	335	12.8	71.1	5372	2	US-09-357-746-1	Sequence 1, Appl
C 263	12.8	71.1	1665	4	US-09-328-352-3670	Sequence 3670, Ap	336	12.8	71.1	5512	4	US-09-866-153-5	Sequence 5, Appl
C 264	12.8	71.1	1707	4	US-09-484-970B-13	Sequence 13, Appl	337	12.8	71.1	5512	4	US-09-693-467A-3	Sequence 5, Appl
C 265	12.8	71.1	1800	4	US-09-248-796A-9868	Sequence 9868, Ap	338	12.8	71.1	5512	4	US-09-270-976-5	Sequence 5, Appl
C 266	12.8	71.1	1830	4	US-09-149-476-129	Sequence 129, App	339	12.8	71.1	5562	4	US-10-204-708-63	Sequence 63, Appl
C 267	12.8	71.1	1879	4	US-09-023-655-742	Sequence 742, App	340	12.8	71.1	5741	1	US-07-706-639-4	Sequence 4, Appl
C 268	12.8	71.1	1929	4	US-09-040-229B-7	Sequence 7, Appl	341	12.8	71.1	5741	1	US-07-998-931-4	Sequence 4, Appl
C 269	12.8	71.1	1971	4	US-09-248-796A-6134	Sequence 6134, Ap	342	12.8	71.1	5868	3	US-08-516-859A-3	Sequence 3, Appl
C 270	12.8	71.1	2025	4	US-09-248-796A-11264	Sequence 11264, A	343	12.8	71.1	5868	4	US-09-586-472-3	Sequence 3, Appl
C 271	12.8	71.1	2045	1	US-08-184-327A-1	Sequence 1, Appl	344	12.8	71.1	5868	4	US-09-528-706-3	Sequence 3, Appl
C 272	12.8	71.1	2046	5	PCT-US95-00670-1	Sequence 1, Appl	345	12.8	71.1	6660	2	US-08-771-602D-1	Sequence 1, Appl
C 273	12.8	71.1	2055	4	US-09-023-655-1036	Sequence 1036, Ap	346	12.8	71.1	6660	3	US-09-232-446B-1	Sequence 1, Appl
C 274	12.8	71.1	2062	1	US-08-073-383-5	Sequence 5, Appl	347	12.8	71.1	7326	1	US-08-194-468-1	Sequence 1, Appl
C 275	12.8	71.1	2062	3	US-08-328-239A-3	Sequence 3, Appl	348	12.8	71.1	7326	3	US-09-514-247A-7	Sequence 7, Appl
C 276	12.8	71.1	2062	5	PCT-US94-06365-5	Sequence 5, Appl	349	12.8	71.1	7326	4	US-09-686-316-1	Sequence 1, Appl
C 277	12.8	71.1	2062	5	PCT-US95-13661-3	Sequence 3, Appl	350	12.8	71.1	7344	3	US-08-961-739-1	Sequence 1, Appl
C 278	12.8	71.1	2115	4	US-09-614-222A-111	Sequence 11, App	351	12.8	71.1	11014	4	US-08-956-711E-91	Sequence 91, Appl
C 279	12.8	71.1	2115	4	US-08-427-993B-2	Sequence 2, Appl	352	12.8	71.1	11014	4	US-08-781-986A-91	Sequence 91, Appl
C 280	12.8	71.1	2333	2	US-08-478-609A-2	Sequence 2, Appl	353	12.8	71.1	11464	3	US-08-991-840A-2	Sequence 2, Appl
C 281	12.8	71.1	2366	4	US-09-270-767-13819	Sequence 13819, A	354	12.8	71.1	11327	1	US-07-906-871-15	Sequence 15, Appl
C 282	12.8	71.1	2414	6	5248599-1	Patent No. 5248599	355	12.8	71.1	19619	4	US-09-596-002-14	Sequence 14, Appl
C 283	12.8	71.1	2429	4	US-09-023-655-75	Sequence 75, Appl	356	12.8	71.1	25603	4	US-09-819-607-3	Sequence 3, Appl
C 284	12.8	71.1	2504	4	US-09-620-312D-778	Sequence 778, App	357	12.8	71.1	38675	4	US-08-311-721A-135	Sequence 135, App
C 285	12.8	71.1	2504	4	US-09-620-312D-778	Sequence 778, App	358	12.8	71.1	44453	3	US-09-146-033-5	Sequence 5, Appl
C 286	12.8	71.1	2652	4	US-09-489-039A-4078	Sequence 4078, Ap	359	12.8	71.1	55827	4	US-09-813-133A-3	Sequence 3, Appl
C 287	12.8	71.1	2753	2	US-07-952-853-5	Sequence 5, Appl	360	12.8	71.1	1230025	4	US-09-198-452A-1	Sequence 1, Appl
C 288	12.8	71.1	2753	2	US-08-914-848-5	Sequence 5, Appl	361	12.8	71.1	1664976	4	US-08-916-621B-1	Sequence 1, Appl
C 289	12.8	71.1	2884	4	US-10-101-464A-291	Sequence 291, App	362	12.8	71.1	1664976	4	US-09-692-570-1	Sequence 1, Appl
C 290	12.8	71.1	2889	1	US-08-537-002A-4	Sequence 4, Appl	363	12.4	68.9	19	4	US-09-696-791-3731	Sequence 371, Ap
C 291	12.8	71.1	2889	3	US-08-863-010-4	Sequence 4, Appl	364	12.4	68.9	21	6	5256770-34	Patent No. 5256770
C 292	12.8	71.1	2889	3	US-09-024-429-4	Sequence 4, Appl	365	12.4	68.9	21	6	5466668-37	Patent No. 5466668
C 293	12.8	71.1	2901	4	US-09-342-325C-43	Sequence 43, Appl	366	12.4	68.9	29	3	US-08-838-151A-42	Sequence 42, Appl
C 294	12.8	71.1	2917	1	US-08-592-936B-20	Sequence 20, Appl	367	12.4	68.9	39	3	US-08-838-151A-22	Sequence 22, Appl
C 295	12.8	71.1	2917	2	US-09-111-573-20	Sequence 20, Appl	368	12.4	68.9	51	4	US-09-443-199C-226	Sequence 226, App
C 296	12.8	71.1	3054	4	US-09-799-451-247	Sequence 247, App	369	12.4	68.9	59	3	US-08-965-762-28	Sequence 28, Appl
C 297	12.8	71.1	3081	4	US-09-248-796A-2025	Sequence 2025, Ap	370	12.4	68.9	59	4	US-09-911-927-28	Sequence 28, Appl
C 298	12.8	71.1	3149	4	US-09-358-856C-2	Sequence 2, Appl	371	12.4	68.9	59	4	US-09-911-882-28	Sequence 28, Appl
C 299	12.8	71.1	3225	4	US-08-658-665-54	Sequence 54, Appl	372	12.4	68.9	87	3	US-08-976-413A-325	Sequence 325, App
C 300	12.8	71.1	3409	2	US-08-658-665-54	Sequence 54, Appl	373	12.4	68.9	168	4	US-09-513-999C-23298	Sequence 23298, A
C 301	12.8	71.1	3409	3	US-08-796-101-18	Sequence 18, Appl	374	12.4	68.9	168	4	US-09-513-999C-10476	Sequence 10476, A
C 302	12.8	71.1	3409	3	US-09-085-273-54	Sequence 54, Appl	375	12.4	68.9	242	4	US-09-513-999C-35084	Sequence 35084, A
C 303	12.8	71.1	3409	3	US-09-916-963-54	Sequence 54, Appl	376	12.4	68.9	258	4	US-09-513-999C-35084	Sequence 35084, A
C 304	12.8	71.1	3499	2	US-08-658-665-52	Sequence 52, Appl	377	12.4	68.9	266	4	US-09-313-294A-1595	Sequence 1595, Ap
C 305	12.8	71.1	3499	2	US-08-796-101-16	Sequence 16, Appl	378	12.4	68.9	267	4	US-09-313-294A-1595	Sequence 1595, Ap
C 306	12.8	71.1	3499	3	US-09-085-273-52	Sequence 52, Appl	379	12.4	68.9	312	4	US-09-710-279-1093	Sequence 1093, Ap
C 307	12.8	71.1	3499	4	US-09-916-963-52	Sequence 52, Appl	380	12.4	68.9	340	4	US-09-270-767-6658	Sequence 8658, Ap
C 308	12.8	71.1	3600	1	US-08-537-002A-5	Sequence 5, Appl	381	12.4	68.9	340	4	US-09-270-767-23940	Sequence 23940, A
C 309	12.8	71.1	3600	3	US-08-863-010-5	Sequence 5, Appl	382	12.4	68.9	398	4	US-09-513-999C-16838	Sequence 16838, A
C 310	12.8	71.1	3600	3	US-09-024-429-5	Sequence 5, Appl	383	12.4	68.9	400	4	US-08-956-111E-2930	Sequence 2930, Ap
C 311	12.8	71.1	3600	4	US-10-101-464A-458	Sequence 458, App	384	12.4	68.9	400	4	US-08-781-986A-2930	Sequence 2930, Ap
C 312	12.8	71.1	3684	4	US-09-205-258-115	Sequence 115, App	385	12.4	68.9	408	3	US-09-134-001C-2068	Sequence 2068, Ap
C 313	12.8	71.1	4338	1	US-08-015-986A-1	Sequence 1, Appl	386	12.4	68.9	450	4	US-09-621-976-1756	Sequence 1756, Ap
C 314	12.8	71.1	4338	2	US-08-446-363-1	Sequence 1, Appl	387	12.4	68.9	532	3	US-08-998-416-145	Sequence 145, App
C 315	12.8	71.1	4396	3	US-09-821-736-1	Sequence 1, Appl	388	12.4	68.9	551	4	US-09-311-021-153	Sequence 153, App
C 316	12.8	71.1	4468	4	US-09-620-312D-243	Sequence 243, App	389	12.4	68.9	564	4	US-09-710-279-2225	Sequence 2225, Ap
C 317	12.8	71.1	4594	2	US-08-658-665-43	Sequence 43, Appl	390	12.4	68.9	594	3	US-08-998-416-143	Sequence 143, App
C 318	12.8	71.1	4594	2	US-08-658-665-45	Sequence 45, Appl	391	12.4	68.9	598	3	US-09-328-111-241	Sequence 241, App
C 319	12.8	71.1	4594	3	US-08-796-101-7	Sequence 7, Appl	392	12.4	68.9	620	4	US-09-270-767-4356	Sequence 4356, App

C 393	12.4	68.9	620	4	US-09-270-767-19638	Sequence 19638, A
C 394	12.4	68.9	623	3	US-09-286-529-9	Sequence 9, Appl1
C 395	12.4	68.9	627	4	US-09-710-279-1403	Sequence 1403, Ap
C 396	12.4	68.9	642	3	US-09-134-001C-1872	Sequence 1872, Ap
C 397	12.4	68.9	683	4	US-10-101-464A-315	Sequence 315, Ap
C 398	12.4	68.9	700	4	US-09-702-705-184	Sequence 184, Ap
C 399	12.4	68.9	700	4	US-09-736-457-184	Sequence 184, Ap
C 400	12.4	68.9	700	4	US-09-614-124B-184	Sequence 184, Ap
C 401	12.4	68.9	700	4	US-09-671-325-184	Sequence 184, Ap
C 402	12.4	68.9	700	4	US-09-580-184-184	Sequence 184, Ap
C 403	12.4	68.9	700	4	US-09-658-824-184	Sequence 184, Ap
C 404	12.4	68.9	724	3	US-08-998-416-627	Sequence 627, Ap
C 405	12.4	68.9	831	2	US-08-743-637B-188	Sequence 188, Ap
C 406	12.4	68.9	885	4	US-09-248-796A-6594	Sequence 6594, Ap
C 407	12.4	68.9	893	3	US-09-286-529-8	Sequence 8, Appl1
C 408	12.4	68.9	1011	3	US-09-095-163-1	Sequence 5, Appl1
C 409	12.4	68.9	1029	3	US-09-485-648-5	Sequence 5, Appl1
C 410	12.4	68.9	1029	4	US-09-503-565-5	Sequence 5, Appl1
C 411	12.4	68.9	1029	4	US-09-485-649-5	Sequence 5, Appl1
C 412	12.4	68.9	1049	4	US-09-270-767-14935	Sequence 14935, A
C 413	12.4	68.9	1073	3	US-08-960-780-37	Sequence 37, Appl
C 414	12.4	68.9	1073	3	US-09-073-898-37	Sequence 37, Appl
C 415	12.4	68.9	1073	4	US-09-073-898-37	Sequence 37, Appl
C 416	12.4	68.9	1104	4	US-09-850-351A-37	Sequence 37, Appl
C 417	12.4	68.9	1104	4	US-09-328-352-877	Sequence 877, Appl
C 418	12.4	68.9	1128	4	US-09-134-000C-1576	Sequence 1576, Ap
C 419	12.4	68.9	1132	4	US-08-956-111E-868	Sequence 868, Ap
C 420	12.4	68.9	1170	3	US-08-781-986A-868	Sequence 868, Ap
C 421	12.4	68.9	1278	3	US-08-467-023-94	Sequence 94, Appl
C 422	12.4	68.9	1284	4	US-09-248-796A-6479	Sequence 6479, Appl
C 423	12.4	68.9	1332	4	US-09-270-767-11937	Sequence 11937, A
C 424	12.4	68.9	1337	3	US-08-467-023-1	Sequence 1, Appl1
C 425	12.4	68.9	1456	4	US-09-056-285A-6	Sequence 6, Appl1
C 426	12.4	68.9	1473	4	US-09-774-528-413	Sequence 413, Ap
C 427	12.4	68.9	1473	4	US-09-056-285A-9	Sequence 9, Appl1
C 428	12.4	68.9	1476	4	US-09-710-279-1797	Sequence 1797, Ap
C 429	12.4	68.9	1514	3	US-09-162-524-4	Sequence 4, Appl1
C 430	12.4	68.9	1542	4	US-09-543-661A-654	Sequence 654, Ap
C 431	12.4	68.9	1566	4	US-09-489-039A-879	Sequence 879, Ap
C 432	12.4	68.9	1612	4	US-09-270-767-14918	Sequence 14918, A
C 433	12.4	68.9	1620	3	US-09-134-001C-2281	Sequence 2281, Ap
C 434	12.4	68.9	1635	4	US-09-248-796A-12936	Sequence 12936, A
C 435	12.4	68.9	1652	4	US-09-976-594-809	Sequence 809, Ap
C 436	12.4	68.9	1658	4	US-09-016-434-1320	Sequence 1320, Ap
C 437	12.4	68.9	1659	4	US-09-248-796A-6048	Sequence 6048, Ap
C 438	12.4	68.9	1664	4	US-09-397-787-68	Sequence 68, Appl
C 439	12.4	68.9	1664	4	US-09-397-787-68	Sequence 68, Appl
C 440	12.4	68.9	1795	4	US-09-799-451-342	Sequence 342, Ap
C 441	12.4	68.9	1827	4	US-09-583-110-1663	Sequence 1663, Ap
C 442	12.4	68.9	1882	3	US-08-501-572-4	Sequence 4, Appl1
C 443	12.4	68.9	1882	3	US-09-040-444-4	Sequence 4, Appl1
C 444	12.4	68.9	1931	4	US-09-429-906B-5	Sequence 5, Appl1
C 445	12.4	68.9	1946	4	US-09-566-921-123	Sequence 123, Ap
C 446	12.4	68.9	1947	3	US-09-134-001C-234	Sequence 234, Ap
C 447	12.4	68.9	2201	3	US-09-330-970-2	Sequence 2, Appl1
C 448	12.4	68.9	2311	4	US-09-774-528-35	Sequence 35, Appl
C 449	12.4	68.9	2403	3	US-09-134-001C-2818	Sequence 2818, Ap
C 450	12.4	68.9	2427	4	US-09-489-039A-4376	Sequence 4376, Ap
C 451	12.4	68.9	2428	4	US-09-799-451-929	Sequence 929, Ap
C 452	12.4	68.9	2447	3	US-08-387-707-12	Sequence 12, Appl
C 453	12.4	68.9	2447	4	US-08-405-271A-12	Sequence 12, Appl
C 454	12.4	68.9	2478	3	US-09-215-694-31	Sequence 31, Appl
C 455	12.4	68.9	2707	4	US-09-870-956-1	Sequence 1, Appl1
C 456	12.4	68.9	2707	4	US-09-870-956-1	Sequence 1, Appl1
C 457	12.4	68.9	2783	4	US-09-620-312D-107	Sequence 55, Appl
C 458	12.4	68.9	2783	4	US-08-887-534A-44	Sequence 107, Appl
C 459	12.4	68.9	2799	4	US-09-527-431-44	Sequence 44, Appl
C 460	12.4	68.9	2799	4	US-09-527-431-44	Sequence 44, Appl
C 461	12.4	68.9	2826	3	US-09-446-861-44	Sequence 44, Appl
C 462	12.4	68.9	2826	3	US-09-010-998-1	Sequence 1, Appl1
C 463	12.4	68.9	2843	4	US-09-620-312D-211	Sequence 211, Ap
C 464	12.4	68.9	3102	4	US-09-522-666-5	Sequence 3682, Ap
C 465	12.4	68.9	3150	3	US-09-710-279-3682	Sequence 3682, Ap
C 466	12.4	68.9	3177	4	US-09-710-279-3717	Sequence 3717, Ap
C 467	12.4	68.9	3247	3	US-08-718-388-4	Sequence 4, Appl1
C 468	12.4	68.9	3320	4	US-09-710-279-3649	Sequence 3649, Ap
C 469	12.4	68.9	3363	3	US-09-252-991A-1356	Sequence 1356, Ap
C 470	12.4	68.9	3767	4	US-09-499-884-11	Sequence 11, Appl
C 471	12.4	68.9	3830	4	US-09-710-279-4213	Sequence 4213, Ap
C 472	12.4	68.9	4079	4	US-09-016-434-1412	Sequence 3773, Ap
C 473	12.4	68.9	4487	4	US-03-484-970B-96	Sequence 1412, Ap
C 474	12.4	68.9	4858	4	US-08-956-171E-379	Sequence 96, Appl
C 475	12.4	68.9	4858	4	US-08-781-986A-379	Sequence 379, Appl
C 476	12.4	68.9	5053	4	US-09-620-312D-590	Sequence 590, Appl
C 477	12.4	68.9	5197	1	US-08-131-365B-53	Sequence 53, Appl
C 478	12.4	68.9	5197	2	US-08-668-123-53	Sequence 53, Appl
C 479	12.4	68.9	5479	4	US-10-164-595-43	Sequence 53, Appl
C 480	12.4	68.9	5735	4	US-09-566-921-12	Sequence 12, Appl
C 481	12.4	68.9	7521	3	US-09-004-838-116	Sequence 116, Appl1
C 482	12.4	68.9	7824	3	US-08-718-388-6	Sequence 6, Appl1
C 483	12.4	68.9	11141	3	US-09-468-671-130	Sequence 10, Appl
C 484	12.4	68.9	11665	4	US-08-961-527-134	Sequence 134, Appl
C 485	12.4	68.9	16388	3	US-08-718-388-8	Sequence 8, Appl1
C 486	12.4	68.9	16592	4	US-08-966-171E-53	Sequence 53, Appl
C 487	12.4	68.9	16592	4	US-08-781-986A-53	Sequence 53, Appl
C 488	12.4	68.9	25603	4	US-09-819-607-3	Sequence 3, Appl1
C 489	12.4	68.9	31328	3	US-09-215-694-19	Sequence 19, Appl
C 490	12.4	68.9	37030	4	US-08-311-731A-25	Sequence 25, Appl
C 491	12.4	68.9	128779	4	US-09-497-855A-38	Sequence 38, Appl
C 492	12.4	68.9	265223	4	US-09-596-002-41	Sequence 41, Appl
C 493	12.4	68.9	53165	4	US-09-214-808-1	Sequence 1, Appl1
C 494	12.2	67.8	19	4	US-09-422-978-9121	Sequence 9121, Ap
C 495	12.2	67.8	20	1	US-08-271-880A-51	Sequence 51, Appl
C 496	12.2	67.8	20	2	US-08-910-408-51	Sequence 51, Appl
C 497	12.2	67.8	20	3	US-09-249-215-51	Sequence 51, Appl
C 498	12.2	67.8	20	4	US-09-544-398B-323	Sequence 323, Appl
C 499	12.2	67.8	28	3	US-09-251-645-19	Sequence 19, Appl
C 500	12.2	67.8	31	4	US-09-603-052-10	Sequence 10, Appl
ALIGNMENTS						
RESULT 1						
US-09-414-436-2/C						
Sequence 2, Application US/09414436						
Patent No. 6294384						
GENERAL INFORMATION:						
APPLICANT: Dell'Acqua, Giorgio						
APPLICANT: Mann, Michael J.						
APPLICANT: Dzuu, Victor J.						
TITLE OF INVENTION: Compositions and Methods Based Upon an Isoform of p53						
FILE REFERENCE: p53f						
CURRENT APPLICATION NUMBER: US/09/414,436						
CURRENT FILING DATE: 1999-10-07						
EARLIER APPLICATION NUMBER: 60/103,849						
EARLIER FILING DATE: 1999-10-09						
NUMBER OF SEQ ID NOS: 5						
SOFTWARE: PatentIn Ver. 2.0						
SEQ ID NO 2						
LENGTH: 1415						
TYPE: DNA						
ORGANISM: Rattus norvegicus						
US-09-414-436-2						
Query Match 100.0%; Score 18; DB 3; Length 1415;						
Best Local Similarity 100.0%; Pred. No. 1.8;						
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps						
Oy 1 CGACTGTCATCTCCAT 18						
Db 41 CGACTGTCATCTCCAT 24						
RESULT 2						
US-08-184-009-214/C						

Sequence 214, Application US/08184009
Patent No. 5933975
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,009
FILING DATE: 19-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066CURTMS
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 1173 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-184-009-214

Query Match 91.1%; Score 16.4; DB 2; Length 1173;
Best Local Similarity 94.4%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18
|||
DB 27 CGACTGTGATCCTCCAT 10

RESULT 3
US-08-458-356-214/c
Sequence 214, Application US/08458356
Patent No. 5942235
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/458,356
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,009
FILING DATE: 19-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066CURTMS
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 1173 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-458-356-214

Query Match 91.1%; Score 16.4; DB 2; Length 1173;
Best Local Similarity 94.4%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18
|||
DB 27 CGACTGTGATCCTCCAT 10

RESULT 4
US-08-796-101-45/c
Sequence 45, Application US/08796101
Patent No. 6183752
GENERAL INFORMATION:
APPLICANT: ERSTEIN, STEPHEN B.
APPLICANT: FINKELE, TOREN
APPLICANT: SPEIR, EDITH
APPLICANT: ZHOU, YI FU
APPLICANT: ZHU, JIANHUI
APPLICANT: ERDILE, LORENE
APPLICANT: PINCUS, STEVEN
TITLE OF INVENTION: RESTENOSIS/ATHEROSCLEROSIS DIAGNOSIS,
NUMBER OF SEQUENCES: 184
TITLE OF INVENTION: PROPHYLAXIS AND THERAPY
CORRESPONDENCE ADDRESS:
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
STREET: 530 FIFTH AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/796,101
FILING DATE: 05-FEB-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: KOMALSKI, THOMAS J.
REGISTRATION NUMBER: 32,147
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 764-5574
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1173 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-796-101-45

Query Match 91.1%; Score 16.4; DB 3; Length 1173;
Best Local Similarity 94.4%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCTCCAT 18
|||||
Db 27 CGACTGTGATCTCTCCAT 10

RESULT 5
US-08-460-736-214/C
Sequence 214, Application US/08460736
Patent No. 6265189

GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,736
FILING DATE: 02-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,009
FILING DATE: 19-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 1173 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-460-736-214

Query Match 91.1%; Score 16.4; DB 3; Length 1173;
Best Local Similarity 94.4%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCTCCAT 18
|||||
Db 27 CGACTGTGATCTCTCCAT 10

RESULT 6
US-09-535-370-214/C
Sequence 214, Application US/09535370
Patent No. 6537594

GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo

Tartaglia, James

Cox, William I.

TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY

NUMBER OF SEQUENCES: 217

CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford

STREET: 530 Fifth Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/535,370

FILING DATE: 24-Mar-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/460,736

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2530

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 214:

SEQUENCE CHARACTERISTICS:

LENGTH: 1173 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 214:

Query Match 91.1%; Score 16.4; DB 4; Length 1173;
Best Local Similarity 94.4%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCTCCAT 18
|||||
Db 27 CGACTGTGATCTCTCCAT 10

RESULT 7
US-09-663-667-214/C
Sequence 214, Application US/09663667
Patent No. 6780407

GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
Tartaglia, James
Cox, William I.

TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/663,667
FILING DATE: 15-Sep-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/184,009
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066CURTMS
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 1173 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-663-667-214

Query Match 91.1%; Score 16.4; DB 4; Length 1173;
Best Local Similarity 94.4%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18
Db 27 CGACTGTGAATCCTCCAT 10

RESULT 8
US-08-585-593A-36/C
Sequence 36, Application US/08585593A
Patent No. 6503706
GENERAL INFORMATION:
APPLICANT: ABKEN, Hinrich J
APPLICANT: ALBERT, Winfried
APPLICANT: JUNGFER, Herbert
TITLE OF INVENTION: METHOD OF IDENTIFYING HUMAN AND ANIMAL
CELLS CAPABLE OF UNLIMITED PROLIFERATION OR TUMOR
TITLE OF INVENTION: FORMATION
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
STREET: 655 Fifteenth Street N.W. Suite 330
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/585,593A
FILING DATE: 16-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02307
FILING DATE: 13-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 43 23 727.4
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Murray, Robert B.
REGISTRATION NUMBER: 22,980
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 638-5000
TELEFAX: (202) 638-4810
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-585-593A-36

Query Match 85.6%; Score 15.4; DB 4; Length 396;
Best Local Similarity 94.1%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCCTCCAT 18
Db 302 GACTGTGAATCCTCCAT 286

RESULT 9
US-09-621-976-1600
Sequence 1600, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jodert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 1600
LENGTH: 527
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: 136..435
NAME/KEY: sig.peptide
LOCATION: 136..279
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 5.80000019073486
OTHER INFORMATION: seq ALTMAGVGVLA/GR
US-09-621-976-1600

Query Match 85.6%; Score 15.4; DB 4; Length 527;
Best Local Similarity 94.1%; Pred. No. 40;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCA 17
Db 430 CGACTGTGAATCCTCCA 446

RESULT 10
US-09-270-767-14686/C
Sequence 14686, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 14686
LENGTH: 2516
TYPE: DNA
ORGANISM: Drosophila melanogaster

US-09-270-767-14686

Query Match

Best Local Similarity 85.6%; Score 15.4; DB 4; Length 2516;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGATCTCTCCA 17
 |||||
 Db 811 CGACTGTGATCTCTCCA 795

RESULT 11

US-09-328-475C-5/c
 ; Sequence 5, Application US/09328475C
 ; Patent No. 6476207

GENERAL INFORMATION:

APPLICANT: Zhang, Jimmy
 APPLICANT: Astel, Jon H.
 APPLICANT: Carroll III, Eddie
 APPLICANT: Endege, Wilson O.
 APPLICANT: Ford, Donna M.
 APPLICANT: Monahan, John E.
 APPLICANT: Schlegel, Robert
 APPLICANT: Steinmann, Kathleen E.
 TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
 FILE REFERENCE: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
 CURRENT APPLICATION NUMBER: US/09/328,475C
 CURRENT FILING DATE: 1999-06-09
 NUMBER OF SEQ ID NOS: 341
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 5
 LENGTH: 1024
 TYPE: DNA
 ORGANISM: Homo Sapien
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)...(1024)
 OTHER INFORMATION: n = A,T,C or G
 US-09-328-475C-5

Query Match

Best Local Similarity 83.3%; Score 15; DB 4; Length 1024;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ACTGTGATCTCTCCAT 18
 |||||
 Db 574 ACTGTGATCTCTCCAT 559

RESULT 12

US-09-248-796A-8906/c
 ; Sequence 8906, Application US/09248796A
 ; Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 FILE REFERENCE: 107196.132
 CURRENT APPLICATION NUMBER: US/09/248,796A
 CURRENT FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: US 60/074,725
 PRIOR FILING DATE: 1998-02-13
 PRIOR APPLICATION NUMBER: US 60/096,409
 PRIOR FILING DATE: 1998-08-13
 NUMBER OF SEQ ID NOS: 28208
 SEQ ID NO 8906
 LENGTH: 1107
 TYPE: DNA
 ORGANISM: Candida albicans
 US-09-248-796A-8906

Query Match

82.2%; Score 14.8; DB 4; Length 1107;

Best Local Similarity 86.9%; Pred. No. 96;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACTGTGATCTCTCCAT 18
 |||||
 Db 283 CTACTGGATCTCTCCAT 266

RESULT 13

US-09-621-976-18989/c
 ; Sequence 18989, Application US/09621976
 ; Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.
 APPLICANT: Jobert, S.Y.
 TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 FILE REFERENCE: GENSET.054PR2
 CURRENT APPLICATION NUMBER: US/09/621,976
 CURRENT FILING DATE: 2000-07-21
 NUMBER OF SEQ ID NOS: 19335
 SOFTWARE: Patent.pm
 SEQ ID NO 18989
 LENGTH: 478
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-621-976-18989

Query Match

Best Local Similarity 80.0%; Score 14.4; DB 4; Length 478;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ACTGTGATCTCTCCAT 18
 |||||
 Db 115 ACTGTGATCTCTCCAT 100

RESULT 14

US-09-186-276B-53/c
 ; Sequence 53, Application US/09186276B
 ; Patent No. 6388173

GENERAL INFORMATION:

APPLICANT: Benfey, Philip
 APPLICANT: Dilaurenzio, Laura
 APPLICANT: Wysocka-Diller, Joanna
 APPLICANT: Malamy, Jocelyn E.
 APPLICANT: Pysh, Leonard
 APPLICANT: Helariutta, Yrjo
 TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses Thereof
 FILE REFERENCE: 5914-075-999
 CURRENT APPLICATION NUMBER: US/09/186,276B
 CURRENT FILING DATE: 1998-11-05
 PRIOR APPLICATION NUMBER: 08/642,445
 PRIOR FILING DATE: 1997-04-24
 PRIOR APPLICATION NUMBER: 08/638,617
 PRIOR FILING DATE: 1996-04-26
 NUMBER OF SEQ ID NOS: 79
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 53
 LENGTH: 1093
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-09-186-276B-53

Query Match

Best Local Similarity 80.0%; Score 14.4; DB 3; Length 1093;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GACTGTGATCTCTCCA 17
 |||||
 Db 164 GACTGTGATCTCTCCA 149

RESULT 15
US-08-842-445-53/C
; Sequence 53, Application US/08842445A
; Patent No. 6441270
; GENERAL INFORMATION:
; APPLICANT: Benfey et al.
; TITLE OF INVENTION: Scarcrow Gene, Promoter and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5914-056-999
; CURRENT APPLICATION NUMBER: US/08/842,445A
; CURRENT FILING DATE: 1997-04-24
; EARLIER APPLICATION NUMBER: 08/638, 617
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 1093
; TYPE: DNA
; ORGANISM: Plant
US-08-842-445-53

Query Match 80.0%; Score 14.4; DB 4; Length 1093;
Best Local Similarity 93.8%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCA 17
Db 164 GACTGTGAATCCTCCA 149

RESULT 16
US-09-186-1888-53/C
; Sequence 53, Application US/09186188B
; Patent No. 6455672
; GENERAL INFORMATION:
; APPLICANT: Benfey et al.
; TITLE OF INVENTION: Scarcrow Gene, Promoter and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5914-074-999
; CURRENT APPLICATION NUMBER: US/09/186,188B
; CURRENT FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 08/842,445
; PRIOR FILING DATE: 1997-04-24
; PRIOR APPLICATION NUMBER: 08/638, 617
; PRIOR FILING DATE: 1996-04-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 1093
; TYPE: DNA
; ORGANISM: Plant
US-09-186-1888-53

Query Match 80.0%; Score 14.4; DB 4; Length 1093;
Best Local Similarity 93.8%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCA 17
Db 164 GACTGTGAATCCTCCA 149

RESULT 17
US-09-919-039-32/C
; Sequence 32, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919, 039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113

PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6727066 3393861CB1
US-09-919-039-32

Query Match 80.0%; Score 14.4; DB 4; Length 1629;
Best Local Similarity 93.8%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACTGTGAATCCTCCAT 18
Db 1074 ACAGTGAATCCTCCAT 1059

RESULT 18
US-08-724-466B-3/C
; Sequence 3, Application US/08724466B
; Patent No. 6063606
; GENERAL INFORMATION:
; APPLICANT: Petkovich, P. Martin, White, Jay A.
; APPLICANT: Beckett, Barbara R., Jones, Glenville
; TITLE OF INVENTION: Retinoid Metabolizing Protein
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Blake, Cassels & Graydon
; STREET: Box 25, Commerce Court West
; CITY: Toronto
; ZIP: M5L 1A9
; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
; COMPUTER: COMPAQ, IBM PC compatible
; OPERATING SYSTEM: MS-DOS 5.1
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,466B
; FILING DATE: October 1, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/667,546
; FILING DATE: June 21, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunt, John C.
; REGISTRATION NUMBER: 36,424
; REFERENCE/DOCKET NUMBER: 50767/00004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 863-4344
; TELEFAX: (416) 863-2653
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1850 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-724-466B-3

Query Match 80.0%; Score 14.4; DB 3; Length 1850;
Best Local Similarity 93.8%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCC 16
Db 1204 CGACTGTGAATCCTCC 1189

RESULT 19
US-08-882-164D-3/C

Sequence 3, Application US/08882164D
Patent No. 6306624
GENERAL INFORMATION:
APPLICANT: Petkovich, P. Martin, White, Jay A.,
APPLICANT: Beckett, Barbara R., Jones, Glenville
TITLE OF INVENTION: Retinoid Metabolizing Protein
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Blake, Cassels & Graydon
STREET: Box 25, Commerce Court West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5L 1A9
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
COMPUTER: COMPAQ, IBM PC compatible
OPERATING SYSTEM: MS-DOS 5.1
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,164D
FILING DATE: June 25, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,546
FILING DATE: June 21, 1996
APPLICATION NUMBER: 08/724,466
FILING DATE: October 1, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunt, John C.
REGISTRATION NUMBER: 36,424
REFERENCE/DOCKET NUMBER: 50767/00010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 863-4344
TELEFAX: (416) 863-2653
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1850 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-882-164D-3

Query Match 80.0%; Score 14.4; DB 3; Length 1850;
Best Local Similarity 93.8%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCC 16
|||||
DB 1204 CGACTGTGAATCCTCC 1189

RESULT 20
US-09-399-081A-7
Sequence 7, Application US/09399081A
Patent No. 6646113
GENERAL INFORMATION:
APPLICANT: DREYFUS, Gideon
APPLICANT: CHARROUX, Bernard
APPLICANT: PELIZZONI, Livio
APPLICANT: FISCHER, Utz
APPLICANT: Liu, Qing
TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS USEFUL FOR THE DIAGNOSIS AND TREAT
FILE REFERENCE: 053893-5027
CURRENT APPLICATION NUMBER: US/09/399,081A
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: US 60/100,866
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patentin version 3.1
SEQ ID NO 7
LENGTH: 3152
TYPE: DNA

ORGANISM: Homo sapiens
US-09-399-081A-7

Query Match 80.0%; Score 14.4; DB 4; Length 3152;
Best Local Similarity 93.8%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCA 17
|||||
DB 2159 GACTGTGAATCCTCCA 2174

RESULT 21
US-08-311-174-4/C
Sequence 4, Application US/08311174
Patent No. 5556776
GENERAL INFORMATION:
APPLICANT: TSUCHIYA, MAKOTO
APPLICANT: MIMA, KIYOSHI
TITLE OF INVENTION: SUCRASE GENE DERIVED FROM CORYNEFORM
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,174
FILING DATE: 23-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 046836/1992
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5556776man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-699-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6911 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-311-174-4

Query Match 80.0%; Score 14.4; DB 1; Length 6911;
Best Local Similarity 93.8%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCA 17
|||||
DB 5815 GACTGTGAATCCTCCA 5800

RESULT 22
US-08-206-176-3/C
Sequence 3, Application US/08206176
Patent No. 5639940
GENERAL INFORMATION:

```

APPLICANT: Garner, Ian
APPLICANT: Dalrymple, Michael A
APPLICANT: Prunkard, Donna E
APPLICANT: Foster, Donald C
TITLE OF INVENTION: Production of Fibrinogen in Transgenic
TITLE OF INVENTION: Animals
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Zymogenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/206,176
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 93-15
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-547-8080 ext 322
TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 8878 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: human fibrinogen B-beta chain
FEATURE:
NAME/KEY: misc RNA
LOCATION: 1..469
FEATURE:
NAME/KEY: exon
LOCATION: 470..583
FEATURE:
NAME/KEY: intron
LOCATION: 584..3257
FEATURE:
NAME/KEY: exon
LOCATION: 3258..3449
FEATURE:
NAME/KEY: intron
LOCATION: 3450..3938
FEATURE:
NAME/KEY: exon
LOCATION: 3939..4122
FEATURE:
NAME/KEY: intron
LOCATION: 4123..5042
FEATURE:
NAME/KEY: exon
LOCATION: 5043..5270
FEATURE:
NAME/KEY: intron
LOCATION: 5271..5830
FEATURE:
NAME/KEY: exon
LOCATION: 5831..5944
FEATURE:
NAME/KEY: intron
LOCATION: 5945..6632
FEATURE:

```

```

NAME/KEY: exon
LOCATION: 6633..6758
FEATURE:
NAME/KEY: intron
LOCATION: 6759..6966
FEATURE:
NAME/KEY: exon
LOCATION: 6967..7252
FEATURE:
NAME/KEY: intron
LOCATION: 7253..7870
FEATURE:
NAME/KEY: exon
LOCATION: 7871..8102
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 8103..8537
FEATURE:
NAME/KEY: misc RNA
LOCATION: 8538..8878
FEATURE:
NAME/KEY: CDS
LOCATION: join(470..583, 3258..3449, 3939..4122, 5043..5270,
5831..5944, 6633..6758, 6967..7252, 7871..8102)
US-08-206-176-3
Query Match 80.0%; Score 14.4; DB 1; Length 8878;
Best Local Similarity 93.8%; Pred. No. 2.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Cy 3 ACTGTGAATCTCTCCAT 18
Db 7090 ACAGTGAATCTCTCCAT 7075
RESULT 23
US-09-614-221A-167/C
Sequence 167, Application US/09614221A
Patent No. 6723837
GENERAL INFORMATION:
APPLICANT: Karunanandaa, Balasubrajini
APPLICANT: Yu, Jaehyuk
APPLICANT: Kishore, Ganesh M.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
WITH STEROL SYNTHESIS AND METABOLISM
FILE REFERENCE: 16516.075
CURRENT APPLICATION NUMBER: US/09/614,221A
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/142,981
PRIOR FILING DATE: 1999-07-12
NUMBER OF SEQ ID NOS: 626
SEQ ID NO 167
LENGTH: 1395
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-167
Query Match 77.8%; Score 14; DB 4; Length 1395;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 5 TGTGAATCTCTCCAT 18
Db 119 TGTGAATCTCTCCAT 106
RESULT 24
US-09-513-999C-36262
Sequence 36262, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclet, A.

```

APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
Patent No. 6783961
FILE REFERENCE: 59, US2, REG
CURRENT APPLICATION NUMBER: US/09/513, 999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 36262
LENGTH: 432
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 381
OTHER INFORMATION: r=a or g
US-09-513-999C-36262

Query Match 76.7%; Score 13.8; DB 4; Length 432;
Best Local Similarity 88.2%; Pred. No. 2.9e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACTGTGAATCCTCCAT 17
Db 65 CCACGTGTAATCCTCCA 81

RESULT 25
US-09-389-681-417/c
Sequence 417, Application US/09389681A
Patent No. 6518237
GENERAL INFORMATION:
APPLICANT: Yuqi, Jjiang
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jjiangchun
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.470C3
CURRENT APPLICATION NUMBER: US/09/389,681A
CURRENT FILING DATE: 1999-09-02
NUMBER OF SEQ ID NOS: 463
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 417
LENGTH: 514
TYPE: DNA
ORGANISM: Homo sapiens
US-09-389-681-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;
Best Local Similarity 88.2%; Pred. No. 3e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
Db 338 GACTGTGAATCCTCCAT 322

RESULT 26
US-09-620-405B-417/c
Sequence 417, Application US/09620405B
Patent No. 6528054
GENERAL INFORMATION:
APPLICANT: Jjiang, Yugu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jjiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

FILE REFERENCE: 210121.470C8
CURRENT APPLICATION NUMBER: US/09/620,405B
CURRENT FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 495
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 417
LENGTH: 514
TYPE: DNA
ORGANISM: Homo sapiens
US-09-620-405B-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;
Best Local Similarity 88.2%; Pred. No. 3e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
Db 338 GACTGTGAATCCTCCAT 322

RESULT 27
US-09-433-826B-417/c
Sequence 417, Application US/09433826B
Patent No. 6579973
GENERAL INFORMATION:
APPLICANT: Jjiang, Yugu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jjiangchun
APPLICANT: Harlocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.470C4
CURRENT APPLICATION NUMBER: US/09/433,826B
CURRENT FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 474
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 417
LENGTH: 514
TYPE: DNA
ORGANISM: Homo sapiens
US-09-433-826B-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;
Best Local Similarity 88.2%; Pred. No. 3e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
Db 338 GACTGTGAATCCTCCAT 322

RESULT 28
US-09-604-287A-417/c
Sequence 417, Application US/09604287A
Patent No. 6586572
GENERAL INFORMATION:
APPLICANT: Jjiang, Yugu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jjiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C7
CURRENT APPLICATION NUMBER: US/09/604,287A
CURRENT FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 489
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 417
LENGTH: 514
TYPE: DNA

ORGANISM: Homo sapiens
US-09-604-287A-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
||| |||||
Db 338 GACTGTGAATCCTCCAT 322

RESULT 29
US-09-834-759-417/C
; Sequence 417, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 417
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-834-759-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
||| |||||
Db 338 GACTGTGAATCCTCCAT 322

RESULT 30
US-09-590-751A-417/C
; Sequence 417, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 417
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-590-751A-417

Query Match 76.7%; Score 13.8; DB 4; Length 514;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
||| |||||
Db 338 GACTGTGAATCCTCCAT 322

RESULT 31
US-09-489-039A-2781
; Sequence 2781, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2781
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2781

Query Match 76.7%; Score 13.8; DB 4; Length 516;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACTGTGAATCCTCCA 17
||| |||||
Db 198 GACTGTGAATCCTCCA 214

RESULT 32
US-09-389-681-258/C
; Sequence 258, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 258
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-389-681-258

Query Match 76.7%; Score 13.8; DB 4; Length 519;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
||| |||||
Db 343 GACTGTGAATCCTCCAT 327

RESULT 33
US-09-620-405B-258/C
; Sequence 258, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.

```
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.470C8
CURRENT APPLICATION NUMBER: US/09/620,405B
CURRENT FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 495
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 258
LENGTH: 519
TYPE: DNA
ORGANISM: Homo sapiens
US-09-620-405B-258
```

```
Query Match          76.7%; Score 13.8; DB 4; Length 519;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 GACTGTGAATCCTCCAT 18
Db      343 GACTTGAATCCTCCAT 327
```

```
RESULT 34
US-09-339-338-258/c
Sequence 258, Application US/09339338A
Patent No. 6573368
GENERAL INFORMATION:
APPLICANT: Yugu, Jiang
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
FILE REFERENCE: 210121.470C2
CURRENT APPLICATION NUMBER: US/09/339,338A
CURRENT FILING DATE: 1999-06-23
NUMBER OF SEQ ID NOS: 315
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 258
LENGTH: 519
TYPE: DNA
ORGANISM: Homo sapiens
US-09-339-338-258
```

```
Query Match          76.7%; Score 13.8; DB 4; Length 519;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 GACTGTGAATCCTCCAT 18
Db      343 GACTTGAATCCTCCAT 327
```

```
RESULT 35
US-09-433-826B-258/c
Sequence 258, Application US/09433826B
Patent No. 6579973
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
FILE REFERENCE: 210121.470C4
CURRENT APPLICATION NUMBER: US/09/433,826B
CURRENT FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 474
SOFTWARE: FastSeq for Windows Version 3.0
```

```
SEQ ID NO 258
LENGTH: 519
TYPE: DNA
ORGANISM: Homo sapiens
US-09-433-826B-258
```

```
Query Match          76.7%; Score 13.8; DB 4; Length 519;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 GACTGTGAATCCTCCAT 18
Db      343 GACTTGAATCCTCCAT 327
```

```
RESULT 36
US-09-604-287A-258/c
Sequence 258, Application US/09604287A
Patent No. 6586572
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.470C7
CURRENT APPLICATION NUMBER: US/09/604,287A
CURRENT FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 489
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 258
LENGTH: 519
TYPE: DNA
ORGANISM: Homo sapiens
US-09-604-287A-258
```

```
Query Match          76.7%; Score 13.8; DB 4; Length 519;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 GACTGTGAATCCTCCAT 18
Db      343 GACTTGAATCCTCCAT 327
```

```
RESULT 37
US-09-834-759-258/c
Sequence 258, Application US/09834759
Patent No. 6680197
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.470C9
CURRENT APPLICATION NUMBER: US/09/834,759
CURRENT FILING DATE: 2001-04-13
NUMBER OF SEQ ID NOS: 547
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 258
LENGTH: 519
TYPE: DNA
ORGANISM: Homo sapiens
US-09-834-759-258
```


Query Match 76.7%; Score 13.8; DB 4; Length 519;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCTCCAT 18
|||
Db 343 GACTGTGAATCTCCAT 327

RESULT 38
US-09-590-751A-258/c
; Sequence 258, Application US/09590751A
; Patent No. 6756477

GENERAL INFORMATION:
; APPLICANT: Yuqin, Jiang
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 258
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-590-751A-258

Query Match 76.7%; Score 13.8; DB 4; Length 519;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCTCCAT 18
|||
Db 343 GACTGTGAATCTCCAT 327

RESULT 39
US-09-270-767-11839/c
; Sequence 11839, Application US/09270767
; Patent No. 6703491
GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11839
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11839

Query Match 76.7%; Score 13.8; DB 4; Length 598;
Best Local Similarity 88.2%; Pred. No. 3.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTCCCA 17
|||
Db 348 CGCAGTGAATCTCCCA 332

RESULT 40
US-09-371-307-79/c
; Sequence 79, Application US/09371307A
; Patent No. 6723897
GENERAL INFORMATION:

APPLICANT: Brown, Sherri M.
APPLICANT: Heck, Gregory R.
APPLICANT: Piller, Kenneth J.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Ellich, Ted D.
APPLICANT: Logusch, Eugene W.
APPLICANT: Rao, Sudabathula
APPLICANT: Ream, Joel E.
APPLICANT: Logusch, Sherry J.
TITLE OF INVENTION: Methods for controlling gibberellin levels
FILE REFERENCE: MOET:216
CURRENT APPLICATION NUMBER: US/09/371,307A
CURRENT FILING DATE: 1999-08-10
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 79
LENGTH: 1041
TYPE: DNA
ORGANISM: Cucurbita maxima
US-09-371-307-79

Query Match 76.7%; Score 13.8; DB 4; Length 1041;
Best Local Similarity 88.2%; Pred. No. 3.3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCTCCAT 18
|||
Db 396 GACTGTGAATCTCCAT 380

RESULT 41
US-09-257-580-1/c
; Sequence 1, Application US/09257580
; Patent No. 6307036
GENERAL INFORMATION:
APPLICANT: Yorkshire Cancer Research
TITLE OF INVENTION: Tumour Suppressor Gene
FILE REFERENCE: Canine p53
CURRENT APPLICATION NUMBER: US/09/257,580
CURRENT FILING DATE: 1999-02-25
PRIOR APPLICATION NUMBER: 9804178.3
PRIOR FILING DATE: 1998-02-28
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1146
TYPE: DNA
ORGANISM: Canis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1146)
US-09-257-580-1

Query Match 76.7%; Score 13.8; DB 3; Length 1146;
Best Local Similarity 88.2%; Pred. No. 3.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCTCCAT 18
|||
Db 17 GACTGTGAATCTCCAT 1

RESULT 42
US-09-257-580-4/c
; Sequence 4, Application US/09257580
; Patent No. 6307036
GENERAL INFORMATION:
APPLICANT: Yorkshire Cancer Research
TITLE OF INVENTION: Tumour Suppressor Gene
FILE REFERENCE: Canine p53
CURRENT APPLICATION NUMBER: US/09/257,580
CURRENT FILING DATE: 1999-02-25
PRIOR APPLICATION NUMBER: 9804178.3

PRIOR FILING DATE: 1998-02-28
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 1146
TYPE: RNA
ORGANISM: canis
US-09-257-580-4

Query Match 76.7%; Score 13.8; DB 3; Length 1146;
Best Local Similarity 88.2%; Pred. No. 3.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
DB 17 GACTGTGAATCCTCCAT 1

RESULT 43
US-09-673-395A-57/c
Sequence 57, Application US/09673395A
Patent No. 6620923
GENERAL INFORMATION:
APPLICANT: SPECHT, THOMAS
APPLICANT: HINZMANN, BERND
APPLICANT: SCHMITT, ARMIN
APPLICANT: PILARSKY, CHRISTIAN
APPLICANT: DAHL, EDGAR
APPLICANT: ROSENTHAL, ANDRE
TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
FILE REFERENCE: ALBRE-12
CURRENT APPLICATION NUMBER: US/09/673,395A
CURRENT FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 637
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 57
LENGTH: 1166
TYPE: DNA
ORGANISM: Homo sapiens
US-09-673-395A-57

Query Match 76.7%; Score 13.8; DB 4; Length 1166;
Best Local Similarity 88.2%; Pred. No. 3.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
DB 1128 GACTGTGAATCCTCCAT 1112

RESULT 44
US-09-270-767-12871/c
Sequence 12871, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12871
LENGTH: 1270
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-12871

Query Match 76.7%; Score 13.8; DB 4; Length 1270;
Best Local Similarity 88.2%; Pred. No. 3.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18

DB 686 GCTGTGACTCTCCAT 670

RESULT 45
US-10-000-489-87/c
Sequence 87, Application US/10000489
Patent No. 6794363
GENERAL INFORMATION:
APPLICANT: Benjamin, Stephane
APPLICANT: Tanaka, Hiroaki
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
FILE REFERENCE: 91.US6.DIV
CURRENT APPLICATION NUMBER: US/10/000,489
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: US 09/924,340
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: PCT/IB01/01715
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: US 60/305,456
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/302,277
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/298,698
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/293,574
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: JPatent
SEQ ID NO 87
LENGTH: 1431
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..151
NAME/KEY: CDS
LOCATION: 152..655
NAME/KEY: 3'UTR
LOCATION: 656..1431
NAME/KEY: polyA_signal
LOCATION: 1399..1404
NAME/KEY: polyA_site
LOCATION: 1416..1431
US-10-000-489-87

Query Match 76.7%; Score 13.8; DB 4; Length 1431;
Best Local Similarity 88.2%; Pred. No. 3.5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
DB 1109 GACTGTGAATCCTCCAT 1093

RESULT 46
US-10-000-489-89/c
Sequence 89, Application US/10000489
Patent No. 6794363
GENERAL INFORMATION:
APPLICANT: Benjamin, Stephane
APPLICANT: Tanaka, Hiroaki
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
FILE REFERENCE: 91.US6.DIV
CURRENT APPLICATION NUMBER: US/10/000,489
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: US 09/924,340
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: PCT/IB01/01715
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: US 60/305,456
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/302,277

PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/298,698
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/293,574
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: JPatent
SEQ ID NO 89
LENGTH: 1431
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..151
NAME/KEY: CDS
LOCATION: 152..655
NAME/KEY: 3'UTR
LOCATION: 656..1431
NAME/KEY: polyA_signal
LOCATION: 1399..1404
NAME/KEY: polyA_site
LOCATION: 1416..1431
US-10-000-489-89

Query Match 76.7%; Score 13.8; DB 4; Length 1431;
Best Local Similarity 88.2%; Pred. No. 3.5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCCTCCAT 18
Db 1109 GACTGTGAATCCTCCAT 1093

RESULT 47
US-09-248-796A-5727
Sequence 5727, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 5727
LENGTH: 1458
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-5727

Query Match 76.7%; Score 13.8; DB 4; Length 1458;
Best Local Similarity 88.2%; Pred. No. 3.5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACTGTGAATCCTCCAT 18
Db 631 GACTGTGAATCCTCCAT 647

RESULT 48
US-09-719-919A-2/c
Sequence 2, Application US/09719919A
Patent No. 6680180
GENERAL INFORMATION:
APPLICANT: Jegia, Timothy James
APPLICANT: ICGen Incorporated
TITLE OF INVENTION: Kv6.2, a Voltage-Gated Potassium Channel Subunit
FILE REFERENCE: 018512-001410US

CURRENT APPLICATION NUMBER: US/09/719,919A
CURRENT FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: US 60/091,466
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: WO PCT/US99/14945
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1518
TYPE: DNA
ORGANISM: Mus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1518)
OTHER INFORMATION: mouse alpha subunit of heteromeric voltage-gated
OTHER INFORMATION: potassium channel Kv6.2
US-09-719-919A-2

Query Match 76.7%; Score 13.8; DB 4; Length 1518;
Best Local Similarity 88.2%; Pred. No. 3.5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCA 17
Db 642 CGACTGTGGCTCTCCA 626

RESULT 49
US-08-687-590-60
Sequence 60, Application US/08687590
Patent No. 6255070
GENERAL INFORMATION:
APPLICANT: Willison, Keith Robert
APPLICANT: Kubota, Hiroshi
APPLICANT: Ashworth, Alan
TITLE OF INVENTION: Folding Proteins
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,590
FILING DATE: 31-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/00192
FILING DATE: 31-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9401791.0
FILING DATE: 31-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9418234.2
FILING DATE: 09-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Baastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 084619-000000US
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 1734 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

```

;      FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..1637
;      US-08-687-590-60

```

Query Match	76.7%;	Score 13.8;	DB 3;	Length 1734;
Best Local Similarity	88.2%;	Pred. No. 3.6e+02;		
Matches 15; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy	1	CGACTGTGAATCCTCCA	17
Db	1576	CCACTGTGATCCTCCA	1592

```

RESULT 50
US-09-673-395A-140
: Sequence 140 Application US/09673395A
: Patent No. 6620923
: GENERAL INFORMATION:
: APPLICANT: SPECHT, THOMAS
: APPLICANT: HINZMANN, BERND
: APPLICANT: SCHMITT, ARMIN
: APPLICANT: PILARSKY, CHRISTIAN
: APPLICANT: DAHL, EDGAR
: APPLICANT: ROSENTHAL, ANDRE
: TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
: FILE REFERENCE: ALBRE-12
: CURRENT APPLICATION NUMBER: US/09/673.395A
: CURRENT FILING DATE: 2000-10-17
: NUMBER OF SEQ ID NOS: 637
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 140
: LENGTH: 1938
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-673-395A-140

```

Query Match	76.7%	Score 13.8	DB 4	Length 1938
Best Local Similarity	88.2%	Pred. No. 3.6e+02		
Matches 15	Conservative 0	Mismatches 2	Indels 0	Gaps 0

```

Qy      2 GACTGTGAATCCTCCAT 18
         |||||
Db      1621 GACTGTGAATCCTTGAT 1637

```

Search completed: January 14, 2005, 17:39:40
Job time : 99.9474 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 14:35:09 ; Search time 2369.37 Seconds
(without alignments)
276.831 Million cell updates/sec

Title: US-09-578-453-1

Perfect score: 18

Sequence: 1 CGACTGGAATCCTCCAT 18

Scoring table: IDENTITY NUC

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST.*

1: gb_est1.*

2: gb_est2.*

3: gb_hic.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gss1.*

9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	237	2	BB571570 BB571570
2	18	100.0	280	2	BB548882 UT-R-A0-a
3	18	100.0	302	2	BB567223 BB567223
4	18	100.0	319	2	BB126959 DEPA0708
5	18	100.0	376	6	CB691142 AMGNNUC:C
6	18	100.0	407	2	BB127536 DEPA1285
7	18	100.0	427	6	CB795162 AMGNNUC:S
8	18	100.0	429	6	CB372427 UT-R-G00-
9	18	100.0	471	6	CB732629 AMGNNUC:M
10	18	100.0	479	2	BB128412 DEPA2162
11	18	100.0	559	2	BB554812 UT-R-E0-C
12	18	100.0	559	6	CB609941 AMGNNUC:U
13	18	100.0	562	7	CP978067 PAR134.R
14	18	100.0	589	6	CB583399 AMGNNUC:N
15	18	100.0	715	7	CB400616 AMGNNUC:R
16	18	100.0	744	7	CB394200 AMGNNUC:R
17	18	100.0	792	7	CB482053 AMGNNUC:R
18	18	100.0	797	7	CB477556 AMGNNUC:R
19	18	100.0	802	7	CB560520 AMGNNUC:R
20	18	100.0	845	6	CB315809 AMGNNUC:R
21	18	100.0	846	6	CB602723 AMGNNUC:R
22	17	94.4	537	5	BB896918 X047G05.P
23	17	94.4	717	6	CB829200 BN40.041G
24	17	94.4	726	8	BB186569 CH230-340

25	17	94.4	1075	9	CNS0591U
26	16.4	91.1	175	7	CK34427 CK34427
27	16.4	91.1	188	7	CB685546 E021008-
28	16.4	91.1	233	2	BB585084 BB585084
29	16.4	91.1	245	2	BB604900 BB604900
30	16.4	91.1	252	7	CB674653 A0951H02-
31	16.4	91.1	266	4	BG062707 L0956G04-
32	16.4	91.1	282	9	CG668669 OST464816
33	16.4	91.1	332	5	BY323607 BY323607
34	16.4	91.1	340	5	BY021968 BY021968
35	16.4	91.1	351	5	BY010085 BY010085
36	16.4	91.1	354	5	BY192197 BY192197
37	16.4	91.1	354	7	D76778 D76778
38	16.4	91.1	397	5	BY043120 BY043120
39	16.4	91.1	407	7	CK577863 CK577863
40	16.4	91.1	428	2	AM209975 AM209975
41	16.4	91.1	430	2	BB846564 BB846564
42	16.4	91.1	447	6	CA895446 B0192B03-
43	16.4	91.1	447	9	CG501482 CG501482
44	16.4	91.1	462	2	BB588888 BB588888
45	16.4	91.1	466	7	CP726263 CP726263
46	16.4	91.1	486	6	CA533945 CA533945
47	16.4	91.1	506	6	CA565317 CA565317
48	16.4	91.1	528	9	CG593379 CG593379
49	16.4	91.1	529	6	CD560905 B0421A08-
50	16.4	91.1	530	6	CA533770 CA533770
51	16.4	91.1	545	7	CP892817 CP892817
52	16.4	91.1	551	9	CG568015 CG568015
53	16.4	91.1	555	6	CF162401 CF162401
54	16.4	91.1	562	4	B1525965 B1525965
55	16.4	91.1	572	5	BQ266772 BQ266772
56	16.4	91.1	575	7	CG683981 CG683981
57	16.4	91.1	585	1	CG508476 CG508476
58	16.4	91.1	596	1	A1731576 A1731576
59	16.4	91.1	600	4	B1985246 B1985246
60	16.4	91.1	608	5	BU923192 7051-51.M
61	16.4	91.1	608	7	CK577814 CK577814
62	16.4	91.1	608	7	CK577851 ISF_WT5_1
63	16.4	91.1	610	7	CP897981 CP897981
64	16.4	91.1	617	1	AA030090 AA030090
65	16.4	91.1	622	2	BB660925 BB660925
66	16.4	91.1	622	6	BY731658 BY731658
67	16.4	91.1	627	2	BE380449 BE380449
68	16.4	91.1	634	4	B1150716 B1150716
69	16.4	91.1	638	2	BB654885 BB654885
70	16.4	91.1	643	2	BB663747 BB663747
71	16.4	91.1	671	2	BB634494 BB634494
72	16.4	91.1	692	2	AM476667 AM476667
73	16.4	91.1	695	4	BJ817162 BJ817162
74	16.4	91.1	703	4	CG963753 CG963753
75	16.4	91.1	713	6	CB948431 CB948431
76	16.4	91.1	745	2	BB916329 BB916329
77	16.4	91.1	758	2	BB581124 BB581124
78	16.4	91.1	764	1	AV751729 AV751729
79	16.4	91.1	768	4	BG864885 BG864885
80	16.4	91.1	775	4	B1105128 B1105128
81	16.4	91.1	785	4	B1652413 B1652413
82	16.4	91.1	821	6	CB56296 CB56296
83	16.4	91.1	846	5	BO953341 BO953341
84	16.4	91.1	872	4	B1416028 B1416028
85	16.4	91.1	881	2	B165854 B165854
86	16.4	91.1	882	6	CB203038 CB203038
87	16.4	91.1	904	5	BU517804 BU517804
88	16.4	91.1	905	5	BU147811 BU147811
89	16.4	91.1	911	2	BF133788 BF133788
90	16.4	91.1	915	5	BO900722 BO900722
91	16.4	91.1	923	5	BO928343 BO928343
92	16.4	91.1	923	5	BO965113 BO965113
93	16.4	91.1	932	5	BO928016 BO928016
94	16.4	91.1	936	4	BG969702 B02837471
95	16.4	91.1	938	4	BG293270 B02390848
96	16.4	91.1	939	6	CB204419 CB204419
97	16.4	91.1	954	2	BB534082 BB534082

AL326667 Tetradon	AL326667
CK34427 H3026D07-	CK34427
CB685546 E021008-	CB685546
BB585084 BB585084	BB585084
BB604900 BB604900	BB604900
CB674653 A0951H02-	CB674653
BG062707 L0956G04-	BG062707
CG668669 OST464816	CG668669
BY323607 BY323607	BY323607
BY021968 BY021968	BY021968
BY010085 BY010085	BY010085
BY192197 BY192197	BY192197
D76778 M074C11 mo	D76778
BY043120 BY043120	BY043120
CK577863 ISF_WT5_1	CK577863
AM209975 U145D12.Y	AM209975
BB846564 BB846564	BB846564
CA895446 B0192B03-	CA895446
CG501482 OST464816	CG501482
BB588888 BB588888	BB588888
CP726263 UT-M-G20-	CP726263
CA533945 C0409B09-	CA533945
CA565317 K0336E11-	CA565317
CG593379 OST50499	CG593379
CD560905 B0421A08-	CD560905
CA533770 C0406D02-	CA533770
CP892817 A0115B04-	CP892817
CG568015 OST194703	CG568015
CF162401 B0711A07-	CF162401
B1525965 602926887	B1525965
BQ266772 NISC_F414	BQ266772
CG683981 E0186D04-	CG683981
CG508476 OST959106	CG508476
A1731576 BNTGH1101	A1731576
B1985246 3133-29.M	B1985246
BU923192 7051-51.M	BU923192
CK577814 ISF_WT5_1	CK577814
CK577851 ISF_WT5_1	CK577851
CP897981 A0232B10-	CP897981
AA030090 M127D03.X	AA030090
BB660925 BB660925	BB660925
BY731658 BY731658	BY731658
BE380449 601270523	BE380449
B1150716 602914801	B1150716
BB654885 BB654885	BB654885
BB663747 BB663747	BB663747
BB634494 BB634494	BB634494
AM476667 uc79h11.Y	AM476667
BJ817162 BJ817162	BJ817162
CG963753 602831395	CG963753
CB948431 AGENCOURT	CB948431
BB916329 601666905	BB916329
BB581124 A02100258	BB581124
AV751729 AV751729	AV751729
BG864885 602799329	BG864885
B1105128 602893382	B1105128
B1652413 603299691	B1652413
CB56296 AGENCOURT	CB56296
BO953341 AGENCOURT	BO953341
B1416028 602987181	B1416028
B165854 601774885	B165854
CB203038 AGENCOURT	CB203038
BU517804 AGENCOURT	BU517804
BU147811 AGENCOURT	BU147811
BF133788 601778707	BF133788
BO900722 AGENCOURT	BO900722
BO928343 AGENCOURT	BO928343
BO965113 AGENCOURT	BO965113
BO928016 AGENCOURT	BO928016
BG969702 602837471	BG969702
BG293270 602390848	BG293270
CB204419 AGENCOURT	CB204419
BB534082 602075070	BB534082

C 98	16.4	91.1	954	5	BUS12162	BUS12162	AGENCOURT	C 171	15.4	85.6	424	2	AM184013	AM184013	xj390f03.x
C 99	16.4	91.1	970	5	B0891359	B0891359	AGENCOURT	C 172	15.4	85.6	427	5	BP624231	BP624231	
C 100	16.4	91.1	980	5	B0958441	B0958441	AGENCOURT	C 173	15.4	85.6	439	4	B1388889	B1388889	BP6247A08
C 101	16.4	91.1	992	5	B0963167	B0963167	AGENCOURT	C 174	15.4	85.6	429	5	BP634296	BP634296	
C 102	16.4	91.1	1053	3	AK079711	AK079711	MUS MUSCU	C 175	15.4	85.6	429	6	CF134938	CF134938	
C 103	16.4	91.1	1093	3	BQ087132	BQ087132	AGENCOURT	C 176	15.4	85.6	433	6	CA043074	CA043074	B8A1MGFB0
C 104	16.4	91.1	1139	2	BF795151	BF795151	MUS MUSCU	C 177	15.4	85.6	443	8	B2857345	B2857345	CH240.232
C 105	16.4	91.1	1159	3	AK086845	AK086845	MUS MUSCU	C 178	15.4	85.6	448	7	H18273	YF48607.r1	
C 106	16.4	91.1	2901	3	AK087834	AK087834	MUS MUSCU	C 179	15.4	85.6	450	8	B2984603	B2984603	PUGH010TD
C 107	16.4	91.1	3276	3	AK083750	AK083750	MUS MUSCU	C 180	15.4	85.6	463	4	BI243070	BI243070	
C 108	16	88.9	171	1	AL610480	AL610480	AGENCOURT	C 181	15.4	85.6	464	7	CN070806	CN070806	1021003E0
C 109	16	88.9	222	8	A2345159	A2345159	LM0079N02	C 182	15.4	85.6	467	4	BI243925	BI243925	RE41743.5
C 110	16	88.9	440	9	CG675515	CG675515	LM0079N02	C 183	15.4	85.6	468	4	BG599937	BG599937	EST504832
C 111	16	88.9	488	1	AL816299	AL816299	AL816299	C 184	15.4	85.6	469	7	CK696304	CK696304	ZFT01-P00
C 112	16	88.9	503	8	AO536049	AO536049	RPCT-11-3	C 185	15.4	85.6	475	1	AV557161	AV557161	AV557161
C 113	16	88.9	551	8	AO143786	AO143786	HS_3075_B	C 186	15.4	85.6	477	5	BQ339673	BQ339673	MK4-NN020
C 114	16	88.9	558	8	CG672956	CG672956	LM0079N02	C 187	15.4	85.6	477	6	CD058843	CD058843	3529_1_10
C 115	16	88.9	563	8	AO536391	AO536391	RPCT-11-3	C 188	15.4	85.6	481	1	AA440442	AA440442	LD15288.5
C 116	16	88.9	1041	8	CC238336	CC238336	CH261-118	C 189	15.4	85.6	485	8	AO247352	AO247352	HS_2059_B
C 117	16	88.9	1057	4	BM923125	BM923125	AGENCOURT	C 190	15.4	85.6	488	5	BO508499	BO508499	EST615914
C 118	16	88.9	1160	8	CC273218	CC273218	CH261-142	C 191	15.4	85.6	489	1	AA846850	AA846850	4342B04.8
C 119	15.4	85.6	148	4	BG983069	BG983069	PM1-CN009	C 192	15.4	85.6	489	1	AA951763	AA951763	
C 120	15.4	85.6	154	7	CN402428	CN402428	LM0079N02	C 193	15.4	85.6	489	4	BU700255	BU700255	
C 121	15.4	85.6	169	2	BF946959	BF946959	MRO-NN019	C 194	15.4	85.6	490	2	BF945442	BF945442	MRO-NN019
C 122	15.4	85.6	202	1	AV211510	AV211510	AV211510	C 195	15.4	85.6	491	2	BF945416	BF945416	
C 123	15.4	85.6	203	7	B8590763	B8590763	B8590763	C 196	15.4	85.6	495	5	BX508749	BX508749	DRF26861
C 124	15.4	85.6	216	2	B8590763	B8590763	B8590763	C 197	15.4	85.6	499	5	BO339220	BO339220	MRO-NN019
C 125	15.4	85.6	224	4	BF958492	BF958492	QV2-NN004	C 198	15.4	85.6	499	9	CR221862	CR221862	Forward
C 126	15.4	85.6	235	2	BF958492	BF958492	QV2-NN004	C 199	15.4	85.6	503	2	BF754997	BF754997	
C 127	15.4	85.6	238	1	AV208003	AV208003	AV208003	C 200	15.4	85.6	506	2	BF409148	BF409148	UI-R-BT1-
C 128	15.4	85.6	240	9	CG192419	CG192419	POTAH86TB	C 201	15.4	85.6	506	2	BF945429	BF945429	MRO-NN019
C 129	15.4	85.6	250	2	BB090142	BB090142	BB090142	C 202	15.4	85.6	509	1	A1774453	A1774453	EST25553
C 130	15.4	85.6	254	1	AV284425	AV284425	AV284425	C 203	15.4	85.6	511	1	A1520014	A1520014	LM399434
C 131	15.4	85.6	255	1	AV215431	AV215431	AV215431	C 204	15.4	85.6	512	7	CN402431	CN402431	
C 132	15.4	85.6	262	1	AV133362	AV133362	AV133362	C 205	15.4	85.6	515	6	CB958705	CB958705	AGENCOURT
C 133	15.4	85.6	286	1	AV168774	AV168774	AV168774	C 206	15.4	85.6	519	2	BE512332	BE512332	946069C06
C 134	15.4	85.6	298	6	CP040507	CP040507	QCT14607	C 207	15.4	85.6	524	7	CO516823	CO516823	633689988
C 135	15.4	85.6	303	4	BG983689	BG983689	PM1-CN015	C 208	15.4	85.6	525	4	BM686378	BM686378	UI-E-CRO-
C 136	15.4	85.6	308	1	AI665168	AI665168	605007F08	C 209	15.4	85.6	527	8	AO404939	AO404939	HS_5048_B
C 137	15.4	85.6	322	2	BE149530	BE149530	RC1-HT025	C 210	15.4	85.6	528	4	BM694111	BM694111	UI-E-CIT-
C 138	15.4	85.6	329	2	AM281052	AM281052	LM0079N02	C 211	15.4	85.6	532	7	CN402426	CN402426	LM006002
C 139	15.4	85.6	332	2	AM281052	AM281052	LM0079N02	C 212	15.4	85.6	532	8	AO697225	AO697225	HS_5519_B
C 140	15.4	85.6	334	7	CG670213	CG670213	OST469229	C 213	15.4	85.6	534	4	BG489781	BG489781	602518847
C 141	15.4	85.6	341	7	CN402422	CN402422	LM004241	C 214	15.4	85.6	540	1	AI519195	AI519195	LD38706.5
C 142	15.4	85.6	345	4	BI396316	BI396316	949047A08	C 215	15.4	85.6	540	4	BM694983	BM694983	UI-E-CIT-
C 143	15.4	85.6	354	4	BI204750	BI204750	EST522790	C 216	15.4	85.6	540	1	A1739892	A1739892	
C 144	15.4	85.6	357	5	BO505461	BO505461	EST612876	C 217	15.4	85.6	541	1	BI775922	BI775922	
C 145	15.4	85.6	360	9	CC794143	CC794143	SALK_0436	C 218	15.4	85.6	547	4	BI679635	BI679635	
C 146	15.4	85.6	361	5	BQ108514	BQ108514	shpID000	C 219	15.4	85.6	548	4	BI214393	BI214393	949011F12
C 147	15.4	85.6	371	5	BY416963	BY416963	BY416963	C 220	15.4	85.6	549	1	AA246762	AA246762	LD05702.5
C 148	15.4	85.6	372	8	AO068980	AO068980	HS_2255_B	C 221	15.4	85.6	549	4	BM822510	BM822510	K-EST0092
C 149	15.4	85.6	373	5	BY403891	BY403891	BY403891	C 222	15.4	85.6	550	1	AI668485	AI668485	605033A03
C 150	15.4	85.6	376	8	AO072605	AO072605	HS_2247_A	C 223	15.4	85.6	550	2	BE509950	BE509950	946069C06
C 151	15.4	85.6	378	4	BI417161	BI417161	949053H12	C 224	15.4	85.6	550	5	BQ513487	BQ513487	EST620902
C 152	15.4	85.6	378	4	BI674193	BI674193	949053H12	C 225	15.4	85.6	551	4	BU556119	BU556119	949053H12
C 153	15.4	85.6	383	1	AL927205	AL927205	AL927205	C 226	15.4	85.6	551	9	CG347657	CG347657	CO1C087TV
C 154	15.4	85.6	383	2	BF804230	BF804230	I15-CI014	C 227	15.4	85.6	554	4	BI167432	BI167432	RE07820.5
C 155	15.4	85.6	385	5	BU667874	BU667874	MC01006E0	C 228	15.4	85.6	559	4	BI396397	BI396397	949063B04
C 156	15.4	85.6	388	6	CA521086	CA521086	949053H12	C 229	15.4	85.6	560	5	BU786357	BU786357	
C 157	15.4	85.6	391	1	AA419137	AA419137	KS11020H1	C 230	15.4	85.6	561	2	BE639387	BE639387	
C 158	15.4	85.6	392	2	BF591538	BF591538	nacl5A808	C 231	15.4	85.6	562	4	BI485698	BI485698	
C 159	15.4	85.6	392	6	BY634653	BY634653	BY634653	C 232	15.4	85.6	564	8	AZ434732	AZ434732	LM0077F19
C 160	15.4	85.6	399	6	BY634653	BY634653	BY634653	C 233	15.4	85.6	565	1	AI533840	AI533840	SD05810.5
C 161	15.4	85.6	399	6	BY668011	BY668011	BY668011	C 234	15.4	85.6	566	2	AM218991	AM218991	EST301463
C 162	15.4	85.6	403	5	BP658184	BP658184	BP658184	C 235	15.4	85.6	566	7	CN402424	CN402424	LM0005325
C 163	15.4	85.6	404	5	BP647985	BP647985	BP647985	C 236	15.4	85.6	566	7	CO169279	CO169279	NDL1_5_HO
C 164	15.4	85.6	405	5	BU667476	BU667476	MC01006E0	C 237	15.4	85.6	569	6	CB070051	CB070051	1625B06.Y
C 165	15.4	85.6	413	1	AV681725	AV681725	AV681725	C 238	15.4	85.6	571	1	AI712140	AI712140	605063G06
C 166	15.4	85.6	417	5	AV402841	AV402841	AV402841	C 239	15.4	85.6	574	1	AI518656	AI518656	LD18286.5
C 167	15.4	85.6	417	6	CF232938	CF232938	CF232938	C 240	15.4	85.6	575	2	BF520360	BF520360	EST457830
C 168	15.4	85.6	419	5	BY395750	BY395750	BY395750	C 241	15.4	85.6	577	4	BM847900	BM847900	K-EST0127
C 169	15.4	85.6	420	8	AZ958129	AZ958129	2M0225114	C 242	15.4	85.6	577	6	CA842369	CA842369	1126D10.Y
C 170	15.4	85.6	420	8	AZ958129	AZ958129	2M0225114	C 243	15.4	85.6	578	1	AI519180	AI519180	LD38672.5

C 244	15.4	85.6	578	6	CD059078	3529	1	10	CD059078	3529	1	10	CD059078	3529	1	10	CD059078	3529	1	10
245	15.4	85.6	578	9	CD171411	OGUDEFSTV			CD171411	OGUDEFSTV			CD171411	OGUDEFSTV			CD171411	OGUDEFSTV		
246	15.4	85.6	584	2	BB667499	601442591			BB667499	601442591			BB667499	601442591			BB667499	601442591		
C 247	15.4	85.6	585	4	B1368953	RE56305.5			B1368953	RE56305.5			B1368953	RE56305.5			B1368953	RE56305.5		
248	15.4	85.6	585	7	CN482270	hw18C10.Y			CN482270	hw18C10.Y			CN482270	hw18C10.Y			CN482270	hw18C10.Y		
249	15.4	85.6	590	4	B1490893	603032132			B1490893	603032132			B1490893	603032132			B1490893	603032132		
C 250	15.4	85.6	591	4	B1233517	949011F12			B1233517	949011F12			B1233517	949011F12			B1233517	949011F12		
251	15.4	85.6	591	2	AM953223	ES7365293			AM953223	ES7365293			AM953223	ES7365293			AM953223	ES7365293		
252	15.4	85.6	591	7	CK004991	AGENCOURT			CK004991	AGENCOURT			CK004991	AGENCOURT			CK004991	AGENCOURT		
253	15.4	85.6	594	2	BF024553	PBP 418 L			BF024553	PBP 418 L			BF024553	PBP 418 L			BF024553	PBP 418 L		
254	15.4	85.6	595	2	BF024230	PVP 593 L			BF024230	PVP 593 L			BF024230	PVP 593 L			BF024230	PVP 593 L		
255	15.4	85.6	595	2	BF024386	PBP 145 L			BF024386	PBP 145 L			BF024386	PBP 145 L			BF024386	PBP 145 L		
256	15.4	85.6	596	8	BZ992625	PUGHT25TD			BZ992625	PUGHT25TD			BZ992625	PUGHT25TD			BZ992625	PUGHT25TD		
C 257	15.4	85.6	597	1	CG761852	LM06B033			CG761852	LM06B033			CG761852	LM06B033			CG761852	LM06B033		
258	15.4	85.6	597	9	AA264286	LD07923.5			AA264286	LD07923.5			AA264286	LD07923.5			AA264286	LD07923.5		
C 259	15.4	85.6	603	5	B0638142	hd19a02.Y			B0638142	hd19a02.Y			B0638142	hd19a02.Y			B0638142	hd19a02.Y		
C 260	15.4	85.6	604	1	AA246917	LD05956.5			AA246917	LD05956.5			AA246917	LD05956.5			AA246917	LD05956.5		
C 261	15.4	85.6	605	4	BM686015	UI-B-CRO-			BM686015	UI-B-CRO-			BM686015	UI-B-CRO-			BM686015	UI-B-CRO-		
C 262	15.4	85.6	606	1	A1456248	LD36135.5			A1456248	LD36135.5			A1456248	LD36135.5			A1456248	LD36135.5		
C 263	15.4	85.6	609	4	B1633351	SD27421.5			B1633351	SD27421.5			B1633351	SD27421.5			B1633351	SD27421.5		
C 264	15.4	85.6	611	2	BE706378	RC1-HWQ25			BE706378	RC1-HWQ25			BE706378	RC1-HWQ25			BE706378	RC1-HWQ25		
C 265	15.4	85.6	611	4	B1213388	RE18775.5			B1213388	RE18775.5			B1213388	RE18775.5			B1213388	RE18775.5		
266	15.4	85.6	611	5	BQ807784	NISC_KK10			BQ807784	NISC_KK10			BQ807784	NISC_KK10			BQ807784	NISC_KK10		
267	15.4	85.6	612	5	AM964616	EST376689			AM964616	EST376689			AM964616	EST376689			AM964616	EST376689		
268	15.4	85.6	613	2	BF024292	PSP 008 L			BF024292	PSP 008 L			BF024292	PSP 008 L			BF024292	PSP 008 L		
C 270	15.4	85.6	613	2	BF024292	PSP 008 L			BF024292	PSP 008 L			BF024292	PSP 008 L			BF024292	PSP 008 L		
C 279	15.4	85.6	632	2	BF024292	PSP 008 L			BF024292	PSP 008 L			BF024292	PSP 008 L			BF024292	PSP 008 L		
C 280	15.4	85.6	632	2	BF024292	PSP 008 L			BF024292	PSP 008 L			BF024292	PSP 008 L			BF024292	PSP 008 L		
C 281	15.4	85.6	639	1	BF024108	PVP 399 L			BF024108	PVP 399 L			BF024108	PVP 399 L			BF024108	PVP 399 L		
C 282	15.4	85.6	640	1	A1295263	LP08845.5			A1295263	LP08845.5			A1295263	LP08845.5			A1295263	LP08845.5		
C 283	15.4	85.6	640	4	B1486287	RE69705.5			B1486287	RE69705.5			B1486287	RE69705.5			B1486287	RE69705.5		
C 284	15.4	85.6	650	5	BQ808015	NISC_KK12			BQ808015	NISC_KK12			BQ808015	NISC_KK12			BQ808015	NISC_KK12		
C 285	15.4	85.6	651	1	A1260669	LP04622.5			A1260669	LP04622.5			A1260669	LP04622.5			A1260669	LP04622.5		
286	15.4	85.6	659	4	B1196759	602755394			B1196759	602755394			B1196759	602755394			B1196759	602755394		
287	15.4	85.6	661	4	BM017318	603643982			BM017318	603643982			BM017318	603643982			BM017318	603643982		
288	15.4	85.6	662	7	CK717642	17526_Swo			CK717642	17526_Swo			CK717642	17526_Swo			CK717642	17526_Swo		
C 289	15.4	85.6	666	6	B1171911	RE13809.5			B1171911	RE13809.5			B1171911	RE13809.5			B1171911	RE13809.5		
290	15.4	85.6	667	4	B1823953	603039135			B1823953	603039135			B1823953	603039135			B1823953	603039135		
291	15.4	85.6	667	9	CG181690	PUIH030TB			CG181690	PUIH030TB			CG181690	PUIH030TB			CG181690	PUIH030TB		
292	15.4	85.6	667	9	CG181690	PUIH030TB			CG181690	PUIH030TB			CG181690	PUIH030TB			CG181690	PUIH030TB		
293	15.4	85.6	669	4	CB113008	K-EST0154			CB113008	K-EST0154			CB113008	K-EST0154			CB113008	K-EST0154		
294	15.4	85.6	670	4	BG773349	602722129			BG773349	602722129			BG773349	602722129			BG773349	602722129		
295	15.4	85.6	670	4	B1561480	603256167			B1561480	603256167			B1561480	603256167			B1561480	603256167		
C 296	15.4	85.6	672	7	W22059	63P5 Human			W22059	63P5 Human			W22059	63P5 Human			W22059	63P5 Human		
297	15.4	85.6	672	9	AG044078	Pan txcg1			AG044078	Pan txcg1			AG044078	Pan txcg1			AG044078	Pan txcg1		
298	15.4	85.6	673	2	BE907298	601500006			BE907298	601500006			BE907298	601500006			BE907298	601500006		
299	15.4	85.6	674	7	CN296314	170004708			CN296314	170004708			CN296314	170004708			CN296314	170004708		
300	15.4	85.6	675	4	BG700653	602682270			BG700653	602682270			BG700653	602682270			BG700653	602682270		
C 301	15.4	85.6	677	6	CA781350	015F06AF			CA781350	015F06AF			CA781350	015F06AF			CA781350	015F06AF		
C 302	15.4	85.6	680	4	BG820014	602782548			BG820014	602782548			BG820014	602782548			BG820014	602782548		
C 303	15.4	85.6	681	8	A2521293	RPCT-11-4			A2521293	RPCT-11-4			A2521293	RPCT-11-4			A2521293	RPCT-11-4		
C 304	15.4	85.6	682	7	CN402429	170005999			CN402429	170005999			CN402429	170005999			CN402429	170005999		
305	15.4	85.6	686	4	BG826767	602751224			BG826767	602751224			BG826767	602751224			BG826767	602751224		
306	15.4	85.6	690	4	B1596327	603243051			B1596327	603243051			B1596327	603243051			B1596327	603243051		
307	15.4	85.6	693	6	CB156629	K-EST015			CB156629	K-EST015			CB156629	K-EST015			CB156629	K-EST015		
308	15.4	85.6	696	7	CG932076	ZMMB0C057			CG932076	ZMMB0C057			CG932076	ZMMB0C057			CG932076	ZMMB0C057		
309	15.4	85.6	697	7	CK591840	esc 1_Yan			CK591840	esc 1_Yan			CK591840	esc 1_Yan			CK591840	esc 1_Yan		
C 310	15.4	85.6	698	4	B1372340	RE59280.5			B1372340	RE59280.5			B1372340	RE59280.5			B1372340	RE59280.5		
311	15.4	85.6	698	4	B1460475	603200873			B1460475	603200873			B1460475	603200873			B1460475	603200873		
312	15.4	85.6	698	6	CD871526	AZ02.118H			CD871526	AZ02.118H			CD871526	AZ02.118H			CD871526	AZ02.118H		
313	15.4	85.6	699	8	BE2021404	oe489d08			BE2021404	oe489d08			BE2021404	oe489d08			BE2021404	oe489d08		
C 314	15.4	85.6	700	8	AG158499	Pan txcg1			AG158499	Pan txcg1			AG158499	Pan txcg1			AG158499	Pan txcg1		
315	15.4	85.6	702	7	CN402420	170005830			CN402420	170005830			CN402420	170005830			CN402420	170005830		
316	15.4	85.6	705	4	B1759405	603043071			B1759405	603043071			B1759405	603043071			B1759405	603043071		
317	15.4	85.6	705	8	BZ021474	oe489d09			BZ021474	oe489d09			BZ021474	oe489d09			BZ021474	oe489d09		
318	15.4	85.6	706	4	BG704777	602688401			BG704777	602688401			BG704777	602688401			BG704777	602688401		
319	15.4	85.6	709	4	BG743670	602632842			BG743670	602632842			BG743670	602632842			BG743670	602632842		
320	15.4	85.6	710	4	B1596935	603242951			B1596935	603242951			B1596935	603242951			B1596935	603242951		
321	15.4	85.6	711	4	B1561401	60255168			B1561401	60255168			B1561401	60255168			B1561401	60255168		
322	15.4	85.6	718	1	AL557717	AL557717			AL557717	AL557717			AL557717	AL557717			AL557717	AL557717		

390	15.4	85.6	869	1	AL536432	AL536432	463	15.4	85.6	1588	3	CR591088	CR591088 full-length
391	15.4	85.6	877	2	BE267279	BE267279	464	15.4	85.6	1592	3	CR599907	CR599907 full-length
392	15.4	85.6	877	2	BM047993	BM047993	465	15.4	85.6	1605	3	CR614623	CR614623 full-length
393	15.4	85.6	878	1	AL522006	AL522006	466	15.4	85.6	1626	3	CR603960	CR603960 full-length
394	15.4	85.6	880	2	BF247221	BF247221	467	15.4	85.6	1642	3	CR609654	CR609654 full-length
395	15.4	85.6	880	6	CD107479	CD107479	468	15.4	85.6	1653	3	CR625078	CR625078 full-length
396	15.4	85.6	881	5	BQ725856	BQ725856	469	15.4	85.6	1656	3	CR620589	CR620589 full-length
397	15.4	85.6	883	4	BI635449	BI635449	470	15.4	85.6	1711	3	CR617488	CR617488 full-length
398	15.4	85.6	885	4	BG537052	BG537052	471	15.4	85.6	1712	3	CR618858	CR618858 full-length
399	15.4	85.6	886	6	CD251096	CD251096	472	15.4	85.6	1712	3	CR618858	CR618858 full-length
400	15.4	85.6	889	6	CD106781	CD106781	473	15.4	85.6	1712	3	CR618858	CR618858 full-length
401	15.4	85.6	890	4	BD518585	BD518585	474	15.4	85.6	1712	3	CR618858	CR618858 full-length
402	15.4	85.6	890	5	BQ857368	BQ857368	475	15.4	85.6	1712	3	CR618858	CR618858 full-length
403	15.4	85.6	892	4	BG675133	BG675133	476	15.4	85.6	1712	3	CR618858	CR618858 full-length
404	15.4	85.6	897	6	CD386873	CD386873	477	15.4	85.6	1712	3	CR618858	CR618858 full-length
405	15.4	85.6	897	7	CNS03787	CNS03787	478	15.4	85.6	1712	3	CR618858	CR618858 full-length
406	15.4	85.6	899	4	BG113617	BG113617	479	15.4	85.6	1712	3	CR618858	CR618858 full-length
407	15.4	85.6	899	4	BG623870	BG623870	480	15.4	85.6	1712	3	CR618858	CR618858 full-length
408	15.4	85.6	899	4	BK324859	BK324859	481	15.4	85.6	1712	3	CR618858	CR618858 full-length
409	15.4	85.6	900	5	BK324859	BK324859	482	15.4	85.6	1712	3	CR618858	CR618858 full-length
410	15.4	85.6	902	5	BU180084	BU180084	483	15.4	85.6	1712	3	CR618858	CR618858 full-length
411	15.4	85.6	903	5	BU904211	BU904211	484	15.4	85.6	1712	3	CR618858	CR618858 full-length
412	15.4	85.6	904	8	BX847125	BX847125	485	15.4	85.6	1712	3	CR618858	CR618858 full-length
413	15.4	85.6	905	2	BZ820113	BZ820113	486	15.4	85.6	1712	3	CR618858	CR618858 full-length
414	15.4	85.6	909	2	BU339399	BU339399	487	15.4	85.6	1712	3	CR618858	CR618858 full-length
415	15.4	85.6	912	4	BG428401	BG428401	488	15.4	85.6	1712	3	CR618858	CR618858 full-length
416	15.4	85.6	914	2	BE619436	BE619436	489	15.4	85.6	1712	3	CR618858	CR618858 full-length
417	15.4	85.6	915	5	BQ942387	BQ942387	490	15.4	85.6	1712	3	CR618858	CR618858 full-length
418	15.4	85.6	917	2	BF529453	BF529453	491	15.4	85.6	1712	3	CR618858	CR618858 full-length
419	15.4	85.6	920	5	BU542600	BU542600	492	15.4	85.6	1712	3	CR618858	CR618858 full-length
420	15.4	85.6	932	2	BF347709	BF347709	493	15.4	85.6	1712	3	CR618858	CR618858 full-length
421	15.4	85.6	932	2	BI602593	BI602593	494	15.4	85.6	1712	3	CR618858	CR618858 full-length
422	15.4	85.6	937	6	CD513871	CD513871	495	15.4	85.6	1712	3	CR618858	CR618858 full-length
423	15.4	85.6	941	4	BG105602	BG105602	496	15.4	85.6	1712	3	CR618858	CR618858 full-length
424	15.4	85.6	944	2	BF024113	BF024113	497	15.4	85.6	1712	3	CR618858	CR618858 full-length
425	15.4	85.6	945	4	BM472787	BM472787	498	15.4	85.6	1712	3	CR618858	CR618858 full-length
426	15.4	85.6	947	8	BZ820107	BZ820107	499	15.4	85.6	1712	3	CR618858	CR618858 full-length
427	15.4	85.6	955	2	BE741394	BE741394	500	15.4	85.6	1712	3	CR618858	CR618858 full-length
428	15.4	85.6	957	2	BG164320	BG164320	501	15.4	85.6	1712	3	CR618858	CR618858 full-length
429	15.4	85.6	966	4	BG115287	BG115287	502	15.4	85.6	1712	3	CR618858	CR618858 full-length
430	15.4	85.6	968	5	BX419451	BX419451	503	15.4	85.6	1712	3	CR618858	CR618858 full-length
431	15.4	85.6	969	4	BG696276	BG696276	504	15.4	85.6	1712	3	CR618858	CR618858 full-length
432	15.4	85.6	971	3	CNS0A1JW	CNS0A1JW	505	15.4	85.6	1712	3	CR618858	CR618858 full-length
433	15.4	85.6	971	7	CO579861	CO579861	506	15.4	85.6	1712	3	CR618858	CR618858 full-length
434	15.4	85.6	983	1	AL539546	AL539546	507	15.4	85.6	1712	3	CR618858	CR618858 full-length
435	15.4	85.6	984	1	AL530745	AL530745	508	15.4	85.6	1712	3	CR618858	CR618858 full-length
436	15.4	85.6	989	9	CNS030LS	CNS030LS	509	15.4	85.6	1712	3	CR618858	CR618858 full-length
437	15.4	85.6	991	5	BX440982	BX440982	510	15.4	85.6	1712	3	CR618858	CR618858 full-length
438	15.4	85.6	995	5	BX398758	BX398758	511	15.4	85.6	1712	3	CR618858	CR618858 full-length
439	15.4	85.6	1009	5	BQ880780	BQ880780	512	15.4	85.6	1712	3	CR618858	CR618858 full-length
440	15.4	85.6	1012	4	BH562959	BH562959	513	15.4	85.6	1712	3	CR618858	CR618858 full-length
441	15.4	85.6	1017	4	BI458505	BI458505	514	15.4	85.6	1712	3	CR618858	CR618858 full-length
442	15.4	85.6	1024	5	BQ068489	BQ068489	515	15.4	85.6	1712	3	CR618858	CR618858 full-length
443	15.4	85.6	1031	6	CD512978	CD512978	516	15.4	85.6	1712	3	CR618858	CR618858 full-length
444	15.4	85.6	1052	4	BM921331	BM921331	517	15.4	85.6	1712	3	CR618858	CR618858 full-length
445	15.4	85.6	1053	5	BQ067722	BQ067722	518	15.4	85.6	1712	3	CR618858	CR618858 full-length
446	15.4	85.6	1060	5	EX384498	EX384498	519	15.4	85.6	1712	3	CR618858	CR618858 full-length
447	15.4	85.6	1062	4	BG164254	BG164254	520	15.4	85.6	1712	3	CR618858	CR618858 full-length
448	15.4	85.6	1064	1	AL541105	AL541105	521	15.4	85.6	1712	3	CR618858	CR618858 full-length
449	15.4	85.6	1072	7	CK029809	CK029809	522	15.4	85.6	1712	3	CR618858	CR618858 full-length
450	15.4	85.6	1075	6	CF272425	CF272425	523	15.4	85.6	1712	3	CR618858	CR618858 full-length
451	15.4	85.6	1091	8	CC274773	CC274773	524	15.4	85.6	1712	3	CR618858	CR618858 full-length
452	15.4	85.6	1158	4	BM473378	BM473378	525	15.4	85.6	1712	3	CR618858	CR618858 full-length
453	15.4	85.6	1269	3	CNS0A0Y4	CNS0A0Y4	526	15.4	85.6	1712	3	CR618858	CR618858 full-length
454	15.4	85.6	1345	3	CR590725	CR590725	527	15.4	85.6	1712	3	CR618858	CR618858 full-length
455	15.4	85.6	1375	3	CNS09YTN	CNS09YTN	528	15.4	85.6	1712	3	CR618858	CR618858 full-length
456	15.4	85.6	1378	8	CC261389	CC261389	529	15.4	85.6	1712	3	CR618858	CR618858 full-length
457	15.4	85.6	1383	3	CNS0A0R1	CNS0A0R1	530	15.4	85.6	1712	3	CR618858	CR618858 full-length
458	15.4	85.6	1386	3	CR621335	CR621335	531	15.4	85.6	1712	3	CR618858	CR618858 full-length
459	15.4	85.6	1440	3	CR602177	CR602177	532	15.4	85.6	1712	3	CR618858	CR618858 full-length
460	15.4	85.6	1516	3	CR609599	CR609599	533	15.4	85.6	1712	3	CR618858	CR618858 full-length
461	15.4	85.6	1561	3	CR613569	CR613569	534	15.4	85.6	1712	3	CR618858	CR618858 full-length
462	15.4	85.6	1564	3	CR601079	CR601079	535	15.4	85.6	1712	3	CR618858	CR618858 full-length

ALIGNMENTS

BB571570 237 bp mRNA linear EST 30-NOV-2000
 BB571570 RIKEN full-length enriched, 10 day neonate skin Mus
 musculus cDNA clone 4733402B01 5', mRNA sequence.

REFERENCE

Alizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T.,
 Carinanci, P., Hanagaki, T., Hayatsu, N., Hiraoka, T., Hirozane, T.,
 Hodojima, Y., Imorani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J.,
 Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Okazaki, Y.,
 Okido, T., Owa, C., Sakai, C., Sakai, K., Sasaki, D., Sato, K.,
 Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
 Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toyota, T.,
 Wachihi, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshiki, A.,
 Muramatsu, M., and Hayashizaki, Y.,
 RIKEN Mouse ESTs (Alizawa, K. et al. 2000)

Unpublished (2000)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan

FEATURES
SOURCE

Source

ORIGIN

ORIGIN

141 CGACTGTGAATCCTCCAT 124

UI-R-A0-af-e-02-0-UI 5', mRNA sequence.
57740003

M Rattus norvegicus

1 (bases 1 to 280)

COMMENT

FEATURES

FEATURES

BOURCE

BOURCE

127 CGACTGTGAATCCTCCAT 110

musculus cDNA clone 2810

Musculus

1 (bases 1 to 302)

Aizawa, K., Akahira, S., Akai

Hodoyama, Y., Imotani, K., I

Nakamura, M., Nishikawa, K., Norioka, T., Ohta, C., Sakai, C.

Suzuki, H., Tagawa, A., Taka

Muramatsu, M. and Hayashizawa, Y. 1990. The structure of the *Arabidopsis thaliana* genome. *Plant Cell* 2: 116-123.

Unpublished (2000)

Db 108 CGACTGTGAATCTCCAT 91

RESULT 6
LOCUS BE127536/c 407 bp mRNA linear EST 15-JUN-2000
DEFINITION DBEAL185 Rat Lambda ZAP Express Library Rattus norvegicus cDNA 5',
mRNA sequence.
ACCESSION BE127536
VERSION BE127536.1 GI:8550273
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 407)
Sleeman,M.A., Morrison,J.G., Strachan,L., Kumble,K.D., Glenn,M.P.,
McGrath,A., Grierson,A., Havukkala,I., Tan,P.L.J. and Watson,J.D.
Expressed sequence tags of cDNA clones from rat dermal papilla
cells.
JOURNAL Unpublished (2000)
COMMENT Contact: Sleeman MA
Biology
Genesis Research and Development Corporation Limited
P.O. Box 50, 1 Fox St, Parnell, Auckland, New Zealand
Tel: 0064 9 373 5600
Fax: 0064 9 373 2189
Email: m.sleeman@genesis.co.nz
Seq primer: T3 forward
High quality sequence stop: 407.
Location/Qualifiers
1..407
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Dark-Agouti"
/db_xref="taxon:10116"
/tissue_type="vibrissae"
/cell_type="dermal papilla"
/clone_lib="Rat Lambda ZAP Express Library"

ORIGIN
Query Match 100.0%; Score 18; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
121 CGACTGTGAATCTCCAT 104

RESULT 7
LOCUS CB795162/c 427 bp mRNA linear EST 16-MAY-2003
DEFINITION AMGNMNC:SRPG2-00011-C10-A srpg2 (10238) Rattus norvegicus cDNA
clone srpg2-00011-cl0 5', mRNA sequence.
ACCESSION CB795162
VERSION CB795162.1 GI:29883639
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 427)
Angen EST Program.
Angen Rat EST Program
Unpublished (2003)
Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00011 row: c column: 10.

FEATURES
source location/Qualifiers
1..427
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone_lib="srpg2-00011-cl0"
/tissue_type="peneal gland brain"
/clone_lib="srpg2 (10238)"
/note="vector: pSPORT1, site_1: SalI, site_2: NotI, peneal
gland brain region"

ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 427;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
121 CGACTGTGAATCTCCAT 104

RESULT 8
LOCUS CD372427/c 429 bp mRNA linear EST 29-MAY-2003
DEFINITION UI-R-G00-csf-b-03-0-UI.r1 UI-R-G00 Rattus norvegicus cDNA clone
UI-R-G00-csf-b-03-0-UI 5', mRNA sequence.
ACCESSION CD372427
VERSION CD372427.1 GI:31156517
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 429)
Bonaldi,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
PUBMED 8889548

JOURNAL
MEDLINE
COMMENT Contact: Soares, MB
Coordinated laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bentto-soares@uiowa.edu
Tissue Procurement: James Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/rat.html>
Seq primer: M13 REVERSE.

FEATURES
source location/Qualifiers
1..429
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone_lib="UI-R-G00-csf-b-03-0-UI"
/tissue_type="Whole embryo"
/dev_stage="embryo 13dpc"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-R-G00"
/note="vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I;
UI-R-G00 is a cDNA library containing the following
tissue(s): rat whole embryo 13dpc. The library was
constructed according to Bonaldi, Lennon and Soares,
Genome Research, 6:791-806, 1996. Denatured RNA was size
fractionated on a 1% agarose gel. First strand cDNA
synthesis was primed with oligo-dT primer containing a Not

I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pTX-Lac vector. The library tag sequence located between the Not I site and the polyA tail is CAGCTCTACT. This library was created for the University of Iowa Program for Rat Gene Discovery and Mapping (Val Sheffield, Bento Soares and Tom Casavant)."

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 429;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18
|||||
Db 239 CGACTGTGAATCCTCCAT 222

RESULT 9
CB732629 471 bp mRNA linear EST 11-APR-2003
LOCUS
DEFINITION AMNCNHC:MBB3-00106-H11-A rat brain E15 (10374) Rattus norvegicus

ACCESSION
CB732629
VERSION
CB732629.1 GI:29799792

KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)

ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 471)
AUTHORS
Amgen EST Program.
TITLE
Amgen Rat EST Program
JOURNAL
Unpublished (2003)
COMMENT
Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00106 row: h column: 11.
Location/Qualifiers

FEATURES
source
1..471

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mbb3-00106-h11"
/tissue_type="brain E15"
/clone_lib="rat Brain E15 (10374)"
/note="Vector: pECB; Site_1: BstXI; Site_2: NotI; rat brain E15"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 471;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18
|||||
Db 200 CGACTGTGAATCCTCCAT 183

RESULT 10
BE128412 479 bp mRNA linear EST 15-JUN-2000
LOCUS
DEFINITION DBPA2162 Rat Lambda ZAP Express Library Rattus norvegicus cDNA 5',
mRNA sequence.

ACCESSION
BE128412
VERSION
BE128412.1 GI:8551165
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 (bases 1 to 479)
AUTHORS
Sleeman,M.A., Morrison,J.G., Strachan,L., Kumble,K.D., Glenn,M.P.,
McGrath,A., Grierison,A., Havukkala,I., Tan,P.L.J. and Watson,J.D.
TITLE
Expressed sequence tags of cDNA clones from rat dermal papilla
cells
JOURNAL
Unpublished (2000)
COMMENT
Contact: Sleeman MA
Biology
Genesis Research and Development Corporation Limited
P.O. Box 50, 1 Fox St, Parnell, Auckland, New Zealand
Tel: 0064 9 373 5600
Fax: 0064 9 373 2189
Email: m.sleeman@genesis.co.nz
Seq primer: T3 forward
High quality sequence stop: 479.
Location/Qualifiers

FEATURES
source
1..479

/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Dark-Agouti"
/db_xref="taxon:10116"
/tissue_type="vibrissae"
/cell_type="dermal papilla"
/clone_lib="Rat Lambda ZAP Express Library"

ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18
|||||
Db 121 CGACTGTGAATCCTCCAT 104

RESULT 11
BF554812 559 bp mRNA linear EST 12-DEC-2000
LOCUS
DEFINITION UI-R-E0-cb-e-01-0-UI-r1 UI-R-E0 Rattus norvegicus cDNA clone

ACCESSION
BF554812
VERSION
BF554812.1 GI:11664584
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)

ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 559)
AUTHORS
Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
PubMed
8889548

COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LIML (info@image.llnl.gov). IMAGE ID= 1770040
Seq primer: M3 Forward.
Location/Qualifiers

FEATURES
source
1..559
/organism="Rattus norvegicus"
/mol_type="mRNA"

/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-E0-cb-e-01-0-UI"
/dev_stage="embryonic"
/lab_host="DH10B (Life Technologies)"
/clone_1lb="UI-R-E0"
/note="Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; This library consists of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture."

ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 559;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18
Db 127 CGACTGGAATCCTCCAT 110

RESULT 12
CB609941/c 559 bp mRNA linear EST 16-MAY-2003
LOCUS AMGNNUC:URTG1-00001-E7-A urtgl (13981) Rattus norvegicus cDNA clone
DEFINITION urtgl-00001-e7 5', mRNA sequence.
ACCESSION CB609941
VERSION CB609941.1 GI:29549590
KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 559)

AUTHORS Amgen EST Program.

TITLE Amgen Rat EST Program

JOURNAL Unpublished (2003)

COMMENT Contact: Dan Fitzpatrick

Amgen, Inc

One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00001 row: e column: 7.

FEATURES Location/Qualifiers

source 1..559

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="urtgl-00001-e7"

/clone_1lb="urtgl (13981)"

/note="Vector: pSPORF1; Rat toxicology library Rearranged from internal pSPORT vector"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 559;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18
Db 113 CGACTGGAATCCTCCAT 96

RESULT 13

CF978067/c

LOCUS CF978067 562 bp mRNA linear EST 24-JUN-2004

DEFINITION FA11354.R Rat retinal ganglion cell Rattus norvegicus cDNA, mRNA

sequence.

ACCESSION CF978067

VERSION CF978067.1 GI:49173525
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 562)

AUTHORS Parkes R.H., Qian J., Goldberg J.L., Quigley H.A. and Zack D.J.

TITLE Gene Expression Profiling of Highly Purified Rat Retinal Ganglion

CELLS Cells

JOURNAL Unpublished (2003)

CONTACT: Parkes RH

Department of Ophthalmology

Johns Hopkins University School of Medicine

600 North Wolfe Street, Baltimore, MD 21287, USA

Tel: 410 502 5230

Fax: 410 502 5382

Email: rfarkes@jhmi.edu.

FEATURES

source

1..562 Location/Qualifiers

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/tissue_type="Retinal Ganglion Cells"

/lab_host="DH10B"

/clone_1lb="Rat retinal ganglion cell"

/note="Organ: Eye; Vector: pDNR-LIB; Site_1: SfiI; Site_2: SfiI; The library was constructed from purified rat

retinal ganglion cells. The Creator SMART cDNA library

method (Clontech) was used. EST analysis was performed on

the unamplified, non-normalized, non-subtracted library."

ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 562;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18
Db 132 CGACTGGAATCCTCCAT 115

RESULT 14

CB583399/c

LOCUS CB583399 589 bp mRNA linear EST 03-APR-2003

DEFINITION AMGNNUC:NRDGI-00174-D10-A nrdgl (10855) Rattus norvegicus cDNA

clone nrdgl-00174-d10 5', mRNA sequence.

ACCESSION CB583399

VERSION CB583399.1 GI:29528865

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 589)

AUTHORS Amgen EST Program.

TITLE Amgen Rat EST Program

JOURNAL Unpublished (2003)

COMMENT Contact: Dan Fitzpatrick

Amgen, Inc

One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00174 row: d column: 10.

FEATURES

source

1..589 Location/Qualifiers

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="nrdgl-00174-d10"

/tissue_type="Dorsal Root Ganglia"

ORIGIN /clone lib="nrcl (10855)"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; rat
dorsal root ganglia"

Query Match 100.0%; Score 18; DB 6; Length 589;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
|||||
DB 103 CGACTGTGAATCCTCCAT 86

RESULT 15
CO400616/c
LOCUS CO400616 715 bp mRNA linear EST 01-JUL-2004
DEFINITION AGENCOURT 27533120 NIH_MGC_254 Rattus norvegicus CDNA clone
IMAGE:7312060 5', mRNA sequence.

ACCESSION CO400616
VERSION CO400616
KEYWORDS EST
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 715)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@db-remail.nih.gov
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
College of Wisconsin
CDNA Library Preparation: Express Genomics
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM15360 row: 0 column: 02
High quality sequence stop: 655.
Location/Qualifiers

FEATURES

1..715
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7312060"
/sex="both"
/tissue_type="Brain - Pooled from several tissues from one
or more individuals"
/lab_host="DH10B TONa"
/clone_lib="NIH_MGC_254"
/note="Organ: brain/CNS; Vector: pExpress-1; Site 1:
ECORV; Site 2: NotI; RNA obtained from brain tissue of 8
wk old animal. Tissues were snap-frozen and kept at -80C
before RNA extraction and purification (TRI-reagent
method). CDNA was primed using oligo-dT primer:
5'-pgACTGATCTGATCGGAGCGCGCCG(T)25-3' and cloned into
the ECORV/NotI sites of pExpress-1. Size-selection >1.25kb
resulted in an average insert size of 2.18 kb. This
primary library is not normalized (normalized library is
NIH_MGC_255) and was constructed by Express Genomics
(Frederick, MD). Note: this is a NIH_MGC library"

ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 715;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
|||||
DB 155 CGACTGTGAATCCTCCAT 138

RESULT 16
CO394200/c
LOCUS CO394200 744 bp mRNA linear EST 01-JUL-2004
DEFINITION AGENCOURT 27533383 NIH_MGC_254 Rattus norvegicus CDNA clone
IMAGE:7311837 5', mRNA sequence.

ACCESSION CO394200
VERSION CO394200
KEYWORDS EST
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 744)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@db-remail.nih.gov
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
College of Wisconsin
CDNA Library Preparation: Express Genomics
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM15360 row: 0 column: 19
High quality sequence stop: 666.
Location/Qualifiers

FEATURES

1..744
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7311837"
/sex="both"
/tissue_type="Brain - Pooled from several tissues from one
or more individuals"
/lab_host="DH10B TONa"
/clone_lib="NIH_MGC_254"
/note="Organ: brain/CNS; Vector: pExpress-1; Site 1:
ECORV; Site 2: NotI; RNA obtained from brain tissue of 8
wk old animal. Tissues were snap-frozen and kept at -80C
before RNA extraction and purification (TRI-reagent
method). CDNA was primed using oligo-dT primer:
5'-pgACTGATCTGATCGGAGCGCGCCG(T)25-3' and cloned into
the ECORV/NotI sites of pExpress-1. Size-selection >1.25kb
resulted in an average insert size of 2.18 kb. This
primary library is not normalized (normalized library is
NIH_MGC_255) and was constructed by Express Genomics
(Frederick, MD). Note: this is a NIH_MGC library"

ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 744;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
|||||
DB 155 CGACTGTGAATCCTCCAT 138

RESULT 17
CK482053/c

```

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7108804"
/tissue_type="kidney, pooled"
/lab_host="DH10B Toba"
/clone_1lb="NIH MGC 235"
/notes="Organ: kidney; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; RNA obtained from pooled kidney tissue from a mix of male and female animals at 8 wk old. Tissues were snap-frozen before RNA extraction and purification (TRI-reagent method). cDNA was primed using oligo-dT primer: 5'-pGACTAGTCTTACATCCGAGCGCCCT(7)-25-' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb resulted in an average insert size of 2.2 kb. This primary library is non-normalized (normalized primary library is NIH MGC 236) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH MGC library."

```

REFERENCE	Rattus.
AUTHORS	1. (bases 1 to 797)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgapbs-rt@mail.nih.gov Tissue Procurement: Howard Jacobs cDNA Library Preparation: Express Genomics cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LLM15013 row: 3 column: 13 High quality sequence stop: 718. Location/Qualifiers 1. .797 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116" /clone="IMAGE:7124271" /issue_type="lung, pooled" /lab_host="DH10B Tona" /clone_id="NIH MGC 230" /note="Organ: lung; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; RNA obtained from pooled lung tissue from a mix of male and female animals at 8 wk old. Tissues were snap-frozen and kept at -80C for two days before RNA extraction and purification (TRI-reagent method). cDNA was primed using oligo-dT primer: 5'-GCAGTGGTCTGAGTCCGAGCGCGCCGCTT25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size selection >1.4kb resulted in an average insert size of 2.3 kb. This primary library is normalized (non-normalized primary library is NIH MGC 231) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH_MGC library."
ORIGIN	
Query Match	100.0%; Score 18; DB 7; Length 797;
Best Local Similarity	100.0%; Pred. No. 91;
Matches	18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY	1 CGACTGTGAATCCTCCAT 18
Db	176 CGACTGTGAATCCTCCAT 159
RESULT 19	
LOCUS	COS60520 802 bp mRNA linear EST 19-JUL-2004
DEFINITION	AGENCOURT_28621896 NIH_MGC_250 Rattus norvegicus cDNA clone
ACCESSION	IMAGE:7384467 5', mRNA sequence.
VERSION	COS60520
KEYWORDS	COS60520.1 GI:50373116
SOURCE	EST.
ORGANISM	Rattus norvegicus (Norway rat)
	Rattus norvegicus
	Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 802)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgapbs-rt@mail.nih.gov

Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical College of Wisconsin
 cDNA Library Preparation: Open Biosystems
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LHAM1549 row: h column: 01
 High quality sequence start: 24
 High quality sequence stop: 725.
 Location/Qualifiers
 1..802

FEATURES

source

/organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7384467"
 /lab_host="DH10B TONA"
 /clone_1lb="NIH MGC 250"
 /note="Organ: thymus; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; RNA obtained from testis tissue of 8 wk old animal. Tissues were snap-frozen and kept at -80C before RNA extraction and purification (TRI-reagent method). cDNA was primed using oligo-dT primer:
 5'-pGACTAGTCTAGATCGAGCGCGCCGCT(7)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb resulted in an average insert size of 1.9 kb. This is a primary library (normalized library is NIH MGC 251) and was constructed by Open Biosystems. Note: This is a NIH MGC library"

ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 802;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTTCAT 18
 |||||
 Db 150 CGACTGTGAATCTTCAT 133

RESULT 20
 CB315809/c 845 bp mRNA linear EST 04-MAR-2003
 LOCUS AGNCCOURT 11525266 NICHED Rr Plt1 Rattus norvegicus cDNA clone
 DEFINITION IMAGE:6888160 5', mRNA sequence.
 CB315809
 ACCESSION CB315809.1 GI:28840044
 VERSION
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 845)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: John C. Marshall, M.D., Ph.D
 cDNA Library Preparation: CLONTECH
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LHCW3143 row: p column: 15
 High quality sequence stop: 390.
 Location/Qualifiers
 1..845
 /organism="Rattus norvegicus"

FEATURES

source

/organism="Rattus norvegicus"

/mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:6888160"
 /tissue_type="Pituitary"
 /lab_host="DH10B"
 /clone_1lb="NICHED Rr Plt1"
 /note="Vector: pDNR-Lib; Site 1: SfiI; Site 2: SfiI; 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCATTTATGCC-3' and 3' adaptor sequence: 5'-ATCTTAGAGCGCGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.23 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 845;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCTTCAT 18
 |||||
 Db 188 CGACTGTGAATCTTCAT 171

RESULT 21
 CK602723/c 846 bp mRNA linear EST 22-JAN-2004
 LOCUS AGNCCOURT 17901070 NIH MGC 234 Rattus norvegicus cDNA clone
 DEFINITION IMAGE:7193583 5', mRNA sequence.
 CK602723
 ACCESSION CK602723.1 GI:41116034
 VERSION
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 846)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LHAM15056 row: b column: 13
 High quality sequence stop: 681.
 Location/Qualifiers
 1..846
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7193583"
 /tissue_type="heart, pooled"
 /lab_host="DH10B TONA"
 /clone_1lb="NIH MGC 234"
 /note="Organ: heart; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; RNA obtained from pooled heart tissue from a mix of male and female animals at 8 wk old. Tissues were snap-frozen and kept at -80C for two days before RNA extraction and purification (TRI-reagent method). cDNA was primed using oligo-dT primer:
 5'-pGACTAGTCTAGATCGAGCGCGCCGCT(7)25-3' and cloned into

FEATURES

source

the ECoRV/NotI sites of pExpress-1. Size-selection >1.4kb resulted in an average insert size of 2.2 kb. This primary library is normalized (non-normalized primary library is NIH_MGC_233) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH_MGC library."

ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 846;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTTCAT 18
|||||
77 CGACTGTGAATCTTCAT 60

RESULT 22
BU896918 537 bp mRNA linear EST 17-OCT-2002
LOCUS X047G05 Populus wood cDNA library Populus tremula x Populus tremuloides cDNA 5 prime, mRNA sequence.
DEFINITION

ACCESSION BU896918 GI:24108125
VERSION BU896918.1
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
Unneberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.
The poplar tree transcriptome: Analysis of expressed sequence tags from multiple libraries

JOURNAL
COMMENT
Unpublished (2002)

Contact: BHALERAO RUPALI R.
Umea Plant Science Center
Department of Plant Physiology
University of Umea, 901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: rupali.bhalerao@plantphys.umu.se.

FEATURES

source
1. .537
/organism="Populus tremula x Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:47664"
/tissue_type="wood"
/clone_lib="Populus wood cDNA library"

ORIGIN

Query Match 94.4%; Score 17; DB 5; Length 537;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTTCAT 17
|||||
434 CGACTGTGAATCTTCAT 418

RESULT 23
CD829200/c 717 bp mRNA linear EST 10-JUL-2003
LOCUS BM40.041G04F011207 BM40 Brassica napus cDNA clone BM40041G04, mRNA
DEFINITION

ACCESSION CD829200
VERSION CD829200.1 GI:32511140
KEYWORDS
SOURCE
ORGANISM

Brassica napus (rape)
Brassicaceae
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 717)

AUTHORS
TITLE
JOURNAL
COMMENT
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante

Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (<http://www.genoplante.com>) and <http://genoplante-info.inbioigen.fr>.

FEATURES

source
1. .717
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jel Neuf"
/db_xref="taxon:3708"
/clone="BM40041G04"
/tissue_type="seed"
/clone_lib="BM40"

ORIGIN

Query Match 94.4%; Score 17; DB 6; Length 717;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTTCAT 18
|||||
650 GACTGTGAATCTTCAT 634

RESULT 24
BZ186569/c 726 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-340J6.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-340J6, genomic survey sequence.
ACCESSION BZ186569
VERSION BZ186569.1 GI:23838462
KEYWORDS
SOURCE
ORGANISM

Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 726)
Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn, A., Gebregeorgis, B., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P., and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment Unpublished (1999)

JOURNAL
COMMENT
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230 (<http://www.chori.org/bacpac/rac230.htm>). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or_ering_information.htm). BAC end page: http://www.tigr.org/cdb/bac_ends/rac/bac_end_intro.html
Plate: 340 row: J column: 6
Seq primer: SP6
Class: BAC ends.

FEATURES

source
1. .726
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHd/MCW"
/db_xref="taxon:10116"
/clone="CH230-340J6"

/sex="Female"
/cell type="Brain"
/clone lib="CHORI-230 Segment 2"
/note="Vector: pTARBA1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SmHsd/MCM) BAC library produced by
Pleier de Jong"

ORIGIN

Query Match 94.4%; Score 17; DB 8; Length 726;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
|||||
305 GACTGTGAATCCTCCAT 289

Db

RESULT 25
CNS0591U 1075 bp DNA linear GSS 01-SEP-2000
LOCUS
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
015B23 of library B from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION AL326667.1 GI:8220256
VERSION
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis

REFERENCE
AUTHORS Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fitzames,C., Mincker,P., Brotier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

TITLE
JOURNAL MEDLINE
PUBMED 20296633
REFERENCE 10835645
AUTHORS
2
Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Fitzames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)

TITLE
JOURNAL MEDLINE
PUBMED 10899143
REFERENCE 3
(bases 1 to 1075)
AUTHORS
JOURNAL
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
Direct Submission
BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr
Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

COMMENT

FEATURES
source
1. 1075
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="015B23"
/clone_lib="B"
/note="Genoscope sequence ID : COAB015CA12C1-end : T7"

ORIGIN

Query Match 94.4%; Score 17; DB 9; Length 1075;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
|||||
198 GACTGTGAATCCTCCAT 214

Db

RESULT 26
CK334427/c 175 bp mRNA linear EST 22-DEC-2003
LOCUS
DEFINITION H3026D07-5 NIA Mouse 15K CDNA Clone Set Mus musculus CDNA clone
H3026D07 5', mRNA sequence.

ACCESSION CK334427
VERSION
KEYWORDS CK334427.1 GI:40290040
EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 175)
Tanaka,T.S., Jaradat,S.A., Lam,M.K., Kargul,G.J., Wang,X.,
Grabovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental CDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

TITLE
JOURNAL MEDLINE
PUBMED 20381348
COMMENT
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lsun.grc.nia.nih.gov
Plate: H3026 row: D column: 07
Seq primer: M13 Reverse
High quality sequence stop: 175
POLYA=No.

FEATURES

source
1. 175
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taeST:H3026D07-5"
/db_xref="taxon:10090"
/clone="H3026D07"
/sex="Clones arrayed from a variety of CDNA libraries"
/dev stage="Clones arrayed from a variety of CDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K CDNA Clone Set"
/note="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; This
clone is among a rearranged set of 15,247 clones from 11
embryo CDNA libraries (including preimplantation stage
embryos from unfertilized egg to blastocyst, embryonic
part of E7.5 embryos, extraembryonic part of E7.5
newborn ovary CDNA library. Average insert size 1.5 kb.
All source libraries are cloned unidirectionally with
Oligo(dT)-Not primers. References include: (1)
Genome-wide expression profiling of mid-gestation
placenta and embryo using a 15,000 mouse developmental
CDNA microarray, 2000. Proc. Natl. Acad. Sci. U.S.A. 97:
9127-9132; (2) Large-scale CDNA analysis reveals phased
gene expression patterns during preimplantation mouse
development, 2000. Development, 127: 1737-1749; (3)
Genome-wide mapping of unselected transcripts from
extraembryonic tissue of 7.5-day mouse embryos reveals
enrichment in the t-complex and under-representation on
the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Query Match 91.1%; Score 16.4; DB 7; Length 175;
Best Local Similarity 94.4%; Pred. No. 5.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGTGATCTCTCAT 18
 Db 123 CGACTGTGATCTCTCAT 106
 RESULT 27 188 bp mRNA linear EST 17-MAY-2004
 CN685546
 LOCUS E0210H08-5 NIA Mouse Embryonic Stem (ES) cell (Lif-, 48 h, high density) cDNA library (Long) Mus musculus cDNA clone NIA:E0210H08
 DEFINITION IMAGE:30850363 5', mRNA sequence.
 CN685546
 ACCESSION CN685546.1 GI:47451992
 VERSION EST
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 188)
 Sharov,A.A., Piao,Y., Maroda,R., Dudekula,D.B., Qian,Y., VanBuren,V., Falco,G., Martin,P.R., Stagg,C.A., Bassez,U.C., Wang,Y., Carter,M.G., Hamatani,T., Alba,K., Akutsu,H., Shatrova,L., Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S., Nagasaja,R., Boheler,K.R., Taub,D., Hodess,R.J., Longo,D.L., Schlesinger,D., Keller,J., Klotz,E., Keleşoğlu,G., Umezawa,A., Vecsö,A.L., Rosenthal,J., Kumath,T., Hogan,B.L., Curci,A., D'Urso,M., Kelson,J., Hide,M. and Ko,M.S.
 Transcription analysis of mouse stem cells and early embryos
 PLOS Biol. 1 (3), 410-419 (2003)
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6620, USA
 Email: chna@igsun.irc.nia.nih.gov
 Seq: E0210 row: H column: 08
 Primer: M13 Reverse
 High quality sequence stop: 188
 POLYA=No.

FEATURES
 source
 1. 188
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129Sv/EvTach"
 /db_xref="nistr:E0210H08-5"
 /db_xref="taxon:10090"
 /clone="NIA:E0210H08 IMAGE:30850363"
 /sex="Male"
 /tissue_type="Embryonic Stem Cell"
 /cell_line="129.3 ES cells"
 /lab_host="DH10B"
 /clone_id="NIA Mouse Embryonic Stem (ES) cell (Lif-, 48 h, high density) cDNA library (Long)"
 /note="vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.irc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544191]). ES plates were plated at density 3x10⁴/cm², on gelatin-coated plates and cultured for 48 hrs at 37 °C, 5% CO₂. Culture medium: DMEM supplemented with 15% FBS, 2 mM L-glutamine, 0.1 mM NEAA, 1mM Sodium pyruvate, 0.1 mM beta-mercaptoethanol, 100 U/ml penicillin, and 100 ug/ml streptomycin. Double-stranded cDNAs were synthesized with 5'-gGACTGTGATCTGATCGGCGCGCCGCTTTTCTTTTCTTTT-3' from 2ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker Lf-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The

products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.4 kb. The library was constructed by Yulan Piao."

ORIGIN
 Query Match 91.1%; Score 16.4; DB 7; Length 188;
 Best Local Similarity 94.4%; Pred. No. 5.7e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGTGATCTCTCAT 18
 Db 145 CGACTGTGATCTCTCAT 128
 RESULT 28 233 bp mRNA linear EST 30-NOV-2000
 BB585084/c
 LOCUS BB585084 RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck Mus musculus cDNA clone
 943003018 5', mRNA sequence.
 BB585084
 ACCESSION BB585084.1 GI:11481628
 VERSION EST
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 233)
 Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P., Hanagaki,T., Hayatsu,N., Hirooka,T., Hirozane,T., Hodoiyama,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Kono,H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T., Watabiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
 Unpublished (2000)
 Contact: Yoshinori Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Sphiro-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoaka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Thermosensitization and thermostabilization of the labile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh,M., Kitsuai,T., Akiyama,U., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Akiyama,Y., Shibata,K., Izawa,M., Kawai,J., Okazaki,Y. and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
 source
 1. 233
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="9430003018"

/tissue_type="embryonic body between diaphragm region and neck"
 /dev_stage="12 days embryo"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck"
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCTCGACGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGACGCTCTTTTCTTTTNNATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI."

Query Match 91.1%; Score 16.4; DB 2; Length 233;
 Best Local Similarity 94.4%; Pred. No. 5.8e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGAGATCTCCAT 18
 |||||
 DB 136 CGACTGAGATCTCCAT 119

RESULT 29
 BB604900/c 245 bp mRNA linear EST 05-DEC-2000
 LOCUS BB604900 RIKEN full-length enriched, 0 day neonate lung Mus
 ACCESSION BB604900
 VERSION BB604900.1 GI:11556302
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)

REFERENCE Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 245)

AUTHORS Alizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P., Hanagaki, T., Hayatsu, N., Hirooka, T., Hirozane, T., Hodojima, Y., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kojima, Y., Kono, H., Kusakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C., Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toya, T., Matabiki, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshiki, A., Yamatsuta, M., and Hayashizaki, Y.
 RIKEN Mouse ESTs (Alizawa, K. et al. 2000)
 Unpublished (2000)
 CONTACT: Yoshinide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/

TITLE
 JOURNAL
 COMMENT
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagakawa, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Thermoregulation and thermoregulation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES

SOURCE

1. 245
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="E03007G03"
 /tissue_type="lung"
 /dev_stage="0 day neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 0 day neonate lung"
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGCACGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGACGCTCTTTTCTTTTNNATCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLX I."

ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 245;
 Best Local Similarity 94.4%; Pred. No. 5.9e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGAGATCTCCAT 18
 |||||
 DB 141 CGACTGAGATCTCCAT 124

RESULT 30
 CN674653/c 252 bp mRNA linear EST 17-MAY-2004
 LOCUS CN674653 A0951H02-5 NIA Mouse Embryonic Stem (ES) cell (11f+, 48 h, high density) cDNA library (Long) Mus musculus cDNA clone NIA:A0951H02
 DEFINITION IMAG:30771733 5', mRNA sequence.
 ACCESSION CN674653
 VERSION CN674653.1 GI:47441104
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)

REFERENCE Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 252)

AUTHORS Sharov, A.A., Piao, Y., Matoba, R., Dudekula, D.B., Qian, Y., VanBuren, V., Falco, G., Martin, P.R., Stagg, C.A., Baasey, U.C., Wang, Y., Carter, M.G., Hamatani, T., Alba, K., Akutsu, H., Sharova, L., Tanaka, T.S., Kimber, W.L., Yoshikawa, T., Jaradat, S.A., Pantano, S., Nagaraia, R., Boheler, K.R., Taub, D., Hodges, R.J., Longo, D.L., Schlusser, D., Keller, J., Klotz, E., Kelsoe, G., Umezawa, A., Vescovi, A.L., Rossant, J., Kumath, T., Hogan, B.L., Curci, A., D'Uso, M., Keisoe, J., Hide, W., and Ko, M.S.
 Transcriptional analysis of mouse stem cells and early embryos
 Proc Biol. 1 (3), 410-419 (2003)
 CONTACT: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@jgun.gsc.nia.nih.gov

TITLE
 JOURNAL
 COMMENT

Plate: A0951 row: H column: 02
Seq primer: M13 Reverse
High quality sequence stop: 252
POLYA=No.

FEATURES

source

Location/Qualifiers

1..252
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129Sv/EvTac"
/db_xref="niaEST:A0951H02-5"
/db_xref="taxon:10090"
/clone="NIA:A0951H02 IMAGE:3077173"
/sex="Male"
/tissue_type="Embryonic Stem Cell"
/cell_line="129.3 ES cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Embryonic Stem (ES) cell (Lif+, 48 h, high density) cDNA library (Long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site.1: SalI; Site.2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]]. ES cells were plated at density 3x10⁴/cm², on gelatin-coated plates and cultured for 48 hrs at 37 OC, 5% CO₂. Culture medium: DMEM supplemented with 15% FBS, 2 mM L-glutamine, 0.1 mM NEAA, 1mM Sodium pyruvate, 0.1 mM beta-mercaptoethanol, 1000 U/ml LIF, 100 U/ml penicillin, and 100 ug/ml streptomycin. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTGCTGCTAGATCGGAGCGGCCGCTTTT-3'] from 2ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker Lp-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.7 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match

Best Local Similarity 94.1%; Score 16.4; DB 7; Length 252;
Pred. No. 5.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18
|||||

Db 127 CGACTGTGACTCCTCCAT 110
|||||

RESULT 31

CG662707/c 266 bp mRNA linear EST 10-JUN-2003
LOCUS
DEFINITION
10956G04-5 NIA Mouse Newborn Kidney cDNA Library2 (Short) Mus
musculus cDNA clone 10956G04 5', mRNA sequence.
ACCESSION
BG062707
VERSION
BG062707.2 GI:31577105
KEYWORDS
EST.

SOURCE
Mus musculus (house mouse)

ORGANISM

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 266)
Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method

JOURNAL

Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE
21429098
PUBMED
11544199

COMMENT

On Jan 25, 2001 this sequence version replaced gi:12533453.
Other_BESTs: L0956G04-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cnaa@lgsun.grc.nia.nih.gov
niaEST (http://lgsun.grc.nia.nih.gov/cDNA/cDNA.html)
Plate: L0956 row: G column: 04
Seq primer: -21M13 Reverse
High quality sequence stop: 266
POLYA=No.

FEATURES

source

Location/Qualifiers

1..266
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:L0956G04-5"
/db_xref="taxon:10090"
/clone="L0956G04"
/tissue_type="Newborn Kidney"
/dev_stage="Newborn"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Kidney cDNA Library2 (Short)"
/note="Vector: pSPORT1 (Invitrogen); Site.1: SalI; Site.2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a short-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]]. In brief, double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTGCTGCTAGATCGGAGCGGCCGCTTTT-3'] from 26 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker Lp-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-L. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 1.5 kb. The library was constructed by Yulan Piao(NIA)."

ORIGIN

Query Match

Best Local Similarity 94.4%; Score 16.4; DB 4; Length 266;
Pred. No. 5.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18
|||||

Db 123 CGACTGTGACTCCTCCAT 106
|||||

RESULT 32

CG668669/c 282 bp DNA linear GSS 02-OCT-2003
LOCUS
DEFINITION
OST7464816 Mus musculus 129Sv/Ev Mus musculus genomic clone
OST7464816, genomic survey sequence.
ACCESSION
CG668669
VERSION
CG668669.1 GI:37492518
KEYWORDS
GSS.

SOURCE

Mus musculus (house mouse)

ORGANISM

REFERENCE
AUTHORS
TITLE
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 282)
 Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, J., Finch, R.A.,
 Piggett, J., Beltranda-Rio, H., Buxton, E.C., Edwards, J., Finch, R.A.,
 Fiddle, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, M., Jang, C.,
 Key, B.W., Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D.,
 Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,
 Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,
 Zhu, O., Person, C. and Sands, A.T.
 Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
 screen to identify potential targets for therapeutic intervention
 Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
 COMMENT Contact: Zambrowicz BP
 OmniBank
 Lexicon Genetics Incorporated
 4000 Research Forest Drive, The Woodlands, TX 77381, USA
 Email: materials@lexgen.com
 Gene trap sequence tag generated by 3' RACE from mouse ES cells as
 described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):608-11)
 Class: Gene Trap.
 FEATURES Location/Qualifiers
 source 1..282
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="129Sv/Ev"
 /db_xref="taxon:10090"
 /clone="OSF464816"
 /cell_type="embryonic stem cell"
 /clone_lib="Mus musculus 129Sv/Ev"
 ORIGIN
 Query Match 91.1%; Score 16.4; DB 9; Length 282;
 Best Local Similarity 94.4%; Pred. No. 6e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGACTGTGATCTCCAT 18
 |||||
 30 CGACTGTGATCTCCAT 13
 |||||
 RESULT 33 332 bp mRNA linear EST 11-DEC-2002
 BY323607/c LOCUS
 DEFINITION Musculus full-length enriched, synovial fibroblasts Mus
 BY323607
 ACCESSION Musculus cDNA clone I03020D04 5', mRNA sequence.
 VERSION BY323607.1 GI:26514011
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 332)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Mikaito, I., Osato, N., Saito, R., Suzuki, H., Yamana, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schinbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schirral, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chaoth, C., Corbani, L.E., Cousins, S., Datta, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmer, S.,
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Kung'u, A.,
 Kurochkin, I.V., Lee, Y., Lemhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Munata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Varvaro, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
 Rogers, U., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 JOURNAL MEDLINE
 PUBLISHED 22354683
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel.: 81-45-503-9222
 Fax: 81-45-503-9226
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ono, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and
 Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome, 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in Riken.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Vasilius Aidinis (Biomedical Sciences
 Research Center 'Al. Fleming' Institute of Immunology 14-16 Al.
 Fleming street 16672 Vatti, Greece) whose assistance we gratefully
 acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 FEATURES Location/Qualifiers
 source 1..332
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="I03020D04"
 /cell_type="synovial fibroblasts"
 /clone_lib="RIKEN full-length enriched, synovial
 fibroblasts"
 ORIGIN
 Query Match 91.1%; Score 16.4; DB 5; Length 332;
 Best Local Similarity 94.4%; Pred. No. 6.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGACTGTGATCTCCAT 18
 |||||
 Db 141 CGACTGTGATCTCCAT 124
 |||||
 RESULT 34 340 bp mRNA linear EST 06-DEC-2002
 BY021968/c LOCUS
 DEFINITION RIKEN full-length enriched, mammary gland RCB-0526
 UY-MC(A) cDNA Mus musculus cDNA clone G830050F09 5', mRNA
 sequence.

FEATURES	
ACCESSION	BY021968
VERSION	BY021968.1
KEYWORDS	GI:26127411
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 340)
	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, T., Bono, H., Kondo, S., Mikaido, I., Oato, N., Saito, R., Suzuki, H., Yamataka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Driegant, T.A., Fleischer, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gierboldt, M., Giesi, C., Godzik, A., Gough, J., Grimmond, S., Gwening, S., Hirokawa, N., Jackson, R.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedziarski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, X., Lemnar, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, D., McKenzia, L., Miki, H., Nagaishima, T., Numata, K., Okido, T., Pavani, W., Perle, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, P.J., Reid, J., Ring, B.Z., Ringwald, M., Santelin, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilmink, L.G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aizawa, T., Fukuda, S., Hara, A., Hashihime, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	22354683
PUBMED	12466851
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Shuhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp URL: http://genome.gsc.riken.jp/ Aizawa, K., Akimura, T., Aikawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Morita, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.go.jp) for further details. Location/Qualifiers

1. 340
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G830050P09"
/tissue_type="mammary gland"
/cell_line="RCB-0526 Jyg-MC(A)"
/clone_lib="RIKEN full-length enriched, mammary gland
RCB-0526 Jyg-MC(A) cDNA"

Query Match 91.1% Score 16.4; DB 5; Length 340;
Best Local Similarity 94.4%; Pred. No. 6.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0.

Db 171 CGACTGTGATCCTCCAT 154

QY 1 CGACTGTGATCCTCCAT 18
|||||
|||||

RESULT 35
BY010085/c

LOCUS
DEFINITION
musculus cDNA clone G73016014 5', mRNA sequence.

ACCESSION
BY010085

VERSION
BY010085.1 GI:26070334

KEYWORDS
EST.

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 351)

REFERENCE
Oikazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gotohori, T., Baldarelli, R., Hill, D. P., Bult, C.,
Hume, D. A., Quackenbush, J., Schmitt, L. M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusa, V.,
Chochia, C., Corradi, L. E., Cousin, S., Dalla, E., Dragani, T. A.,
Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Guelincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,
Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Oikido, T., Pavan, W. J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J. T., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
Sanderlin, A., Schneider, C., Sample, C. A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L. G., Wyrshaw-Boris, A., Yangisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayashu, N., Hitozane-Kishikawa, T., Kono, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Yoshino, M., Waterston, R., Lander, E. S.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
MEDLINE
PUBMED
12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Shinto-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

SOURCE

Location/Qualifiers
1. .351
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G730016014"
/tissue_type="lung"
/cell_line="RCB-0558 LLC"
/clone_lib="RIKEN full-length enriched, lung RCB-0558 LLC cDNA"

ORIGIN

Query Match 91.1%; Score 16.4; DB 5; Length 351;
Best Local Similarity 94.4%; Pred. No. 6.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CGACTGTGAATCTCCAT 18
172 CGACTGTGAATCTCCAT 155

RESULT 36
BY192197/c
LOCUS
DEFINITION
BY192197 354 bp mRNA linear EST 10-DEC-2002
BY192197 RIKEN full-length enriched, NOD-derived CD11c +ve dendritic cells Mus musculus cDNA clone F630326B04 5', mRNA sequence.
VERSION
BY192197
KEYWORDS
BY192197.1 GI:26368080
EST.
Mus musculus (house mouse)
ORIGIN

REFERENCE
AUTHORS
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 354)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Mikalido, L., Otsu, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Balderelli, R., Hill, D. P., Bull, C., Hume, D. A., Quackenbush, J., Schraml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gerlold, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Guinichich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pette, G., Pesole, G.,

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

SOURCE

Location/Qualifiers
1. .354
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="taxon:10090"
/clone="F630326B04"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched, NOD-derived CD11c +ve dendritic cells"

ORIGIN

Query Match 91.1%; Score 16.4; DB 5; Length 354;
Best Local Similarity 94.4%; Pred. No. 6.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CGACTGTGAATCTCCAT 18
149 CGACTGTGAATCTCCAT 132

RESULT 37
LOCUS D76778/c 354 bp mRNA linear EST 07-OCT-1996
DEFINITION MUS74C11 mouse embryonal carcinoma cell line F9 Mus musculus cDNA
clone 74C11, mRNA sequence.
ACCESSION D76778
VERSION D76778.1 GI:1596448
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 354)
AUTHORS Nishiguchi, S., Sakuma, R., Nomura, M., Zou, Z., Jearamakulwong, J.,
Jou, T., Yasunaga, T. and Shimada, K.
TITLE A catalogue of genes in mouse embryonal carcinoma F9 cells
identified with expressed sequence tags
JOURNAL J. Biochem. 119 (4), 749-767 (1996)
MEDLINE 96337530
PubMed 8743579
COMMENT Contact: Kazunori Shimada
Department of Medical Genetics, Division of Molecular Biomedicine
Research Institute for Microbial Diseases, Osaka University
3-1, Yamadaoka, Suita, Osaka, 565, Japan
Tel: 06-879-8325
Fax: 06-879-8326
FEATURES
source
1. 354
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="74C11"
/clone_lib="mouse embryonal carcinoma cell line F9"
ORIGIN
Query Match 91.1%; Score 16.4; DB 7; Length 354;
Best Local Similarity 94.4%; Pred. No. 6.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATCCTCCAT 18
|||||
106 CGACTGTGATCCTCCAT 89
|||||
RESULT 38
LOCUS BY043120/c 397 bp mRNA linear EST 06-DEC-2002
DEFINITION BY043120 RIKEN full-length enriched, pooled tissues, cell_line=TRB-55B88, etc. Mus musculus cDNA clone 1730027D23 5',
mRNA sequence.
ACCESSION BY043120
VERSION BY043120.1 GI:26148563
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 397)
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Satoh, R., Suzuki, H., Yamana, T.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Haesegawa, T., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
Balasov, S., Beisel, K. W., Blake, J. A., Bradt, D., Bruscia, V.,
Chochua, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Guelinckich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Kongaya, A.,
Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
Maltais, L., Marchionni, L., McKenzie, L., Mikti, H., Nagashima, T.,

Nunata, K., Okido, T., Pavan, W. J., Perrea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramchandran, S.,
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
Verardo, R., Wagner, L., Walestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L. G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavoian, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Aizawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shiba, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 12466851
PubMed 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa, K., Akimura, T., Aizawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
FEATURES
source
1. 397
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="1730027D23"
/clone_lib="RIKEN full-length enriched, pooled tissues,
cell_line=TRB-55B88, etc."
/note="Pooled tissues: (sex=mix, cell_line=TRB-55B88),
(sex=mix, cell_line=CTL-1722 L5178Y-R)"
ORIGIN
Query Match 91.1%; Score 16.4; DB 5; Length 397;
Best Local Similarity 94.4%; Pred. No. 6.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATCCTCCAT 18
|||||
Db 150 CGACTGTGATCCTCCAT 133
|||||
RESULT 39

CK577863/c
 LOCUS CK577863 407 bp mRNA linear EST 16-JAN-2004
 DEFINITION 1st WTS_18124 AD-wrmCDNA library Caenorhabditis elegans cDNA 5' similar to F38A5.3, mRNA sequence.
 ACCESSION CK577863
 VERSION CK577863.1 GI:40961531
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditidae; Rhabditidae; Pelodierinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 407)
 Li, S., Armstrong, C.M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P.O., Han, J.D., Chesneau, A., Hao, T., Goldberg, D.S., Li, N., Martineau, M., Raul, J.F., Lamesch, P., Xu, L., Tewari, M., Wong, S.L., Zhang, L.V., Bertiz, G.F., Jacotot, J., Vaglio, P., Reboul, O., Hirozane-Kishikawa, T., Li, Q., Gabel, H.W., Elvira, A., Baumgartner, B., Rose, D.J., Yu, H., Bosak, S., Segueria, R., Fraser, A., Mango, S.E., Saxton, W.M., Strom, S., Van Den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K.C., Harper, J.W., Cusick, M.E., Roth, F.P., Hill, D.E. and Vidal, M.
 A Map of the Interactome Network of the Metazoan C. elegans Science (2004) in press
 CONTACT: Vidal M
 Dana Farber Cancer Institute
 Marc Vidal Laboratory
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739
 Email: Marc.Vidal@dfci.harvard.edu
 For the purpose of protein interaction mapping, we generated a C. elegans cDNA library (AD-wrmCDNA) in which poly(dT)-primed reverse transcribed cDNA are fused to the AD-encoding sequence of the yeast transcription factor GAL4. cDNAs were generated and cloned into the two hybrid vector pPC86 This interacting Sequence Tag 1st WTS_18124 (F38A5.3) interacts as a prey with the bait F08G2.5
 PCR primers
 FORWARD: CGCGTTGGAATCACTACAGG
 BACKWARD: CGAGACTGACCAACTCTGGG
 Insert Length: 407 Std Error: 47.00
 Plate: 227 Row: 02 Column: F
 Seq primer: CGCGTTGGAATCACTACAGG
 High quality sequence stop: 406
 POLYA=No.
 FEATURES
 SOURCE location/Qualifiers
 1..407
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /sex="male, hermaphrodite"
 /dev_stage="embryo, L1, L2, L3, L4, adult, dauer"
 /clone_lib="AD-wrmCDNA library"
 /note="Vector: pPC86; For the purpose of protein interaction mapping, we generated a C. elegans cDNA library (AD-wrmCDNA) in which poly(dT)-primed reverse transcribed cDNA are fused to the AD-encoding sequence of the yeast transcription factor GAL4. This library was made with poly(A) + RNA isolated from mated populations of wild-type (N2 strain) animals of all stages of development including embryonic, larval (L1 to L4 stages), adults and dauer. Approximately equal quantities of RNA from different populations were acquired. cDNAs were generated and cloned into the two hybrid vector pPC86. The library contains ~3*10⁶ clones. Reference - GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFsomes - Walhout AJ, Temple GF, Brasch MA, Hartley JL, Lironi MA, van den Heuvel S, Vidal M - Methods Enzymol. 2000;328:575-92"

Best Local Similarity 94.4%; Pred. No. 6.3e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CGACTGGAATCTCCAT 18
 Db 101 CGACTGGAATCTCCCT 84
 RESULT 40
 AM209975/c 428 bp mRNA linear EST 03-DEC-1999
 LOCUS AM209975
 DEFINITION u149d12.y1 Rasbass mouse MOV 9 5 optic vesicle Mus musculus cDNA clone IMAGE:2101655 5' similar to gb:X01237 Mouse mRNA for cellular tumour antigen p53 (MOUSE);, mRNA sequence.
 ACCESSION AM209975
 VERSION AM209975.1 GI:6515915
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 428)
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Stepcoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 CONTACT: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:998587
 Seq primer: Primer name ambiguous.
 FEATURES
 SOURCE location/Qualifiers
 1..428
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CD-1"
 /db_xref="taxon:10090"
 /tissue_type="optic vesicle and lens placode"
 /dev_stage="embryo - 9.5 dpc"
 /lab_host="DH10B"
 /clone_lib="Rasbass mouse MOV 9 5 optic vesicle"
 /note="Organ: eye; Vector: pSPORT1; mRNA made from developing eye tissue, cDNA made by oligo-dT priming with NotI oligo. Salt adaptor (5'-TCAGCCAGCGCG-3') ligated to 5' ends. Size-selected with cDNA size fractionation resin, average insert size 1.3 kb. Primary library, non-amplified. Library constructed by Dr. Pen Rasbass (penr@gu.mrc.ac.uk)."
 ORIGIN
 Query Match 91.1%; Score 16.4; DB 2; Length 428;
 Best Local Similarity 94.4%; Pred. No. 6.3e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CGACTGGAATCTCCAT 18
 Db 125 CGACTGGAATCTCCAT 108
 RESULT 41
 BB846564/c 430 bp mRNA linear EST 26-NOV-2001
 LOCUS BB846564
 DEFINITION BB846564 RIKEN full-length enriched, adult male kidney Mus musculus cDNA clone F53003D19 5', mRNA sequence.

Query Match

91.1%; Score 16.4; DB 7; Length 407;

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
B8846564	B8846564.1	GI:17084939							
EST.									
Mus musculus									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus									
1 (bases 1 to 430)									
Akimura, T., Arikawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayata, N., Hiramoto, K., Hiroaka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Komori, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numaaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, K., Sakai, K., Sakazume, N., Saeki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takenashi, F., Takaku-Akihira, S., Tanaka, T., Tomaru, A., Toya, T., Watabiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.									
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)									
Unpublished (2001)									
Contact: Yoshihide Hayashizaki									
Laboratory for Genome Exploration Research Group, RIKEN Genomic									
Sciences Center (GSC), Yokohama Institute									
The Institute of Physical and Chemical Research (RIKEN)									
1-7-22 Saito-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan									
Tel.: 81-45-503-9222									
Fax: 81-45-503-9216									
Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/									
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,									
Itoh, M., Komori, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.									
Normalization and subtraction of cap-trapper-selected cDNAs to									
prepare full-length cDNA libraries for rapid discovery of new									
genes. Genome Res. 10 (10), 1617-1630 (2000)									
wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,									
Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,									
Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.									
and Hayashizaki, Y.									
RIKEN integrated sequence analysis (RISA) system--384-format									
sequencing pipeline with 384 multichannel sequencer. Genome Res.									
10 (11), 1757-1771 (2000)									
Komori, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,									
Sugahara, Y. and Hayashizaki, Y.									
Computer-based methods for the mouse full-length cDNA									
encyclopedia: real-time sequence clustering for construction of a									
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)									
Please visit our web site (http://genome.gsc.riken.go.jp) for									
further details.									
e mouse tissues.									
Location/Qualifiers									
1. 430									
/organism="Mus musculus"									
/mol_type="mRNA"									
/db_xref="taxon:10090"									
/clone="F530003D19"									
/sex="male"									
/tissue_type="kidney"									
/dev_stage="adult"									
/lab_host="SOLR"									
/clone_lib="RIKEN full-length enriched, adult male kidney									
/note="Site 1: XhoI, Site 2: SctI, cDNA library was									
prepared and sequenced in Mouse Genome Encyclopedia									
Project of Genome Exploration Research Group in Riken									
Genomic Sciences Center and Genome Science Laboratory in									
RIKEN. Division of Experimental Animal Research in Riken									
contributed to prepare mouse tissues. 1st strand cDNA was									
primed with a primer [5'									
GAGGAGAGAGCGGCGCCGACCTGAGTTTCTTTTCTTTT 3'] cDNA was									
prepared by using triethanol thermo-activated reverse									
transcriptase and subsequently enriched for full-length by									
cap-trapper. Second strand cDNA was prepared with the									
primer adapter of sequence [5'									
GAGGAGAGAGAGATCCCAAGGCTCAATTATTTATTAACCCCCCCCC 3']									
cDNA was cleaved with XhoI and SctI. "									

ORIGIN	Query Match	91.1%;	Score 16.4;	DB 2;	Length 430;
	Best Local Similarity	94.4%;	Pred. No. 6.3e+02;		
	Matches 17;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1 CGACTGTGAATCCTCCAT	18			
Db	141 CGACTGTACTCTTCAT	124			
RESULT 42	CA895446	447 bp	mRNA	linear	EST 20-DEC-2002
LOCUS	B0192B03-5N NIA Mouse Neural Stem Cell (Differentiated) cDNA				
DEFINITION	Library (Long) Mus musculus cDNA clone NIA:B0192B03 IMAGE:30102638				
ACCESSION	5', mRNA sequence.				
VERSION	CA895446				
KEYWORDS	CA895446.1 GI:27346995				
ORGANISM	EST.				
SOURCE	Mus musculus (house mouse)				
	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 447)				
AUTHORS	Piao, X., Dudekula, D. B., Qian, Y., Martin, P. R., Alpa, K., Vescovi, A. L.				
	and Ko, M. S. H.				
TITLE	Systematic Analyses of NIA Mouse Neural Stem Cell (Differentiated)				
JOURNAL	cDNA Library (Long)				
COMMENT	Unpublished (2002)				
	Contact: Dawood B. Dudekula				
	Laboratory of Genetics				
	National Institute on Aging/National Institutes of Health				
	333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA				
	Email: cdna@gsun.grc.nia.nih.gov				
	Plate: B0192 row: B column: 03				
	Seq primer: -21M13 Reverse				
	High quality sequence stop: 447				
	POLYA=No.				
FEATURES	location/Qualifiers				
source	1..447				
	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/strain="CD1"				
	/db_xref="labST:B0192B03-5N"				
	/db_xref="taxon:10090"				
	/clone="NIA:B0192B03 IMAGE:30102638"				
	/dev_stage="Adult"				
	/lab_host="DH10B"				
	/clone_id="NIA Mouse Neural Stem Cell (Differentiated)				
	cDNA Library (Long)"				
	/note="Vector: pSPORI1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://gsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). PMID: 115441991). Total RNAs were obtained from Dr. Angelo L. Vescovi (Institute for Stem Cell Research, Italy). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGCGAGCCGCCCTTTTCTTTT-3'] from 2.0 Microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker lR-Sal4, purified by phenol/chloroform, and separated from free linkers by Geniorm 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Geniorm 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORI1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about				

ORIGIN 3.2 kb. The library was constructed by Yulan Piao."

Query Match 91.1%; Score 16.4; DB 6; Length 447;
Best Local Similarity 94.4%; Pred. No. 6.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 114 CGACTGTGATCCTCCAT 97

RESULT 43
CG501482/c 447 bp DNA linear GSS 01-OCT-2003
LOCUS OST45353 Mus musculus 1295v/Er Mus musculus genomic clone OST45353.
DEFINITION genomic survey sequence.

ACCESSION CG501482.1 GI:37275240

VERSION CG501482.1 GI:37275240

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 447)

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE Zambrowicz, B.P., Abuln, A., Ramirez-Solis, R., Richter, L.J.,

JOURNAL 4000 Research Forest Drive, The Woodlands, TX 77381, USA

COMMENT Email: materials@lexgen.com

FEATURES Location/Qualifiers

SOURCE 1. .447

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="1295v/Er"

/db_xref="taxon:10090"

/clone="OST45353"

/cell_type="embryonic stem cell"

/clone_idb="Mus musculus 1295v/Er"

ORIGIN

Query Match 91.1%; Score 16.4; DB 9; Length 447;

Best Local Similarity 94.4%; Pred. No. 6.4e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCCTCCAT 18

Db 34 CGACTGTGATCCTCCAT 17

RESULT 44

BB858888 RIKEN full-length enriched, Nullipotent stem cell CRU-2070

DEFINITION NE cDNA Mus musculus cDNA clone G43003B20 5', mRNA sequence.

ACCESSION BB858888.1 GI:17100342

VERSION BB858888.1 GI:17100342

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 462)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

QY 1 CGACTGTGATCCTCCAT 18

Db 144 CGACTGTGATCCTCCAT 127

RESULT 45

CF726263/c 466 bp mRNA linear EST 09-OCT-2003

LOCUS UI-M-GZO-cjm-d-20-0-UI-r1 NIH_BMP_GZO Mus musculus cDNA clone

DEFINITION IMAGE:30603259 5', mRNA sequence.

ACCESSION CF726263.1 GI:37600431

VERSION CF726263.1 GI:37600431

KEYWORDS

SOURCE

FEATURES Location/Qualifiers

SOURCE 1. .462

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C3H/J"

/db_xref="taxon:10090"

/clone="G43003B20"

/cell_type="Nullipotent stem cell"

/cell_line="CRU-2070 NE"

/dev_stage="7 days embryo"

/clone_idb="RIKEN full-length enriched, Nullipotent stem

cell CRU-2070 NE cDNA"

ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 462;

Best Local Similarity 94.4%; Pred. No. 6.4e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

KEYWORDS
SOURCE EST.
ORGANISM Mus musculus (house mouse)

REFERENCE
 1 (bases 1 to 466)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE
 JOURNAL
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
source
 Location/Qualifiers
 1..466
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30603259"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP G20"
 /note="Organ: Eye; Vector: pYX-Aac; Site 1: EcoR I; Site 2: Not I. The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to RNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Aac vector. The library tag sequence located between the Not I site and the polyA tail is TTTATGAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
 Query Match 91.1%; Score 16.4; DB 7; Length 466;
 Best Local Similarity 94.4%; Pred. No. 6.4e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY
 1 CGACTGTGATCTCCAT 18
 |||||
 |||||

Db
 231 CGACTGTGATCTCCAT 214

RESULT 46
CAS33945/c 486 bp mRNA linear EST 18-NOV-2002
LOCUS C0409809-5N NIA Mouse ES Cell (LIF-) cDNA library (long) Mus
DEFINITION musculus cDNA clone NIA:C0409809 IMAGE:30010484 5', mRNA sequence.
ACCESSION CAS333945
VERSION CAS333945.1 GI:25065602
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 486)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradit, S.A., Boheler, K.R. and Ko, M.S.H.
TITLE Systematic Analyses of NIA Mouse ES Cell (LIF-) cDNA library (long)

JOURNAL
COMMENT Unpublished (2001)
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: C0409 row: B column: 09
 Seg primer: M13 Reverse
 High quality sequence stop: 486
 POLYA=No.

FEATURES
source
 Location/Qualifiers
 1..486
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129/Sv x 129/Sv-CP"
 /db_xref="t1aEST:C0409809-5N"
 /db_xref="taxon:10090"
 /clone="NIA:C0409809 IMAGE:30010484"
 /tissue_type="ES Cell (LIF-)"
 /cell_line="F1 ES cells"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse ES Cell (LIF-) cDNA library (long)"
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cdna>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). PMID: 11544193). Total RNAs were obtained from Dr. Kenneth R. Boheler (National Institute on Aging, USA). ES cells were cultured without feeder cells in the absence of LIF for 4 hours or 18 hours. Equimolar mixture of these RNA samples was used for the library construction. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen: 5'-pGACTGTCTGATCGGACGCCGCCCTTTTCTTTT-3') from 21 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker IL-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.4 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN
 Query Match 91.1%; Score 16.4; DB 6; Length 486;
 Best Local Similarity 94.4%; Pred. No. 6.4e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY
 1 CGACTGTGATCTCCAT 18
 |||||
 |||||

Db
 121 CGACTGTGATCTCCAT 104

RESULT 47
CAS65317/c 506 bp mRNA linear EST 19-NOV-2002
LOCUS K0336E11-5N NIA Mouse Osteoblast cDNA library (long) Mus musculus
DEFINITION cDNA clone NIA:K0336E11 IMAGE:30057754 5', mRNA sequence.
ACCESSION CAS65317
VERSION CAS65317.1 GI:25109988
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 506)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,

/clone="NIA:B0421A08 IMAGE:30440647"
/tissue_type="E6.5 whole embryo"
/dev_stage="whole embryo including extraembryonic tissues
at 6.5-days postcoitum"
/lab_host="DH10B"
/clone_lib="NIA Mouse E6.5 Whole Embryo cDNA Library
(long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were extracted from a pool of 7 embryos at 6.5-days
postcoitum. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer (Invitrogen):
5'-GGACTGATCTAGATCGGAGCGCGCCCTTTT-3'
from 0.53 ug of total RNA, treated with T4 DNA polymerase,
and purified by ethanol-precipitation. The cDNAs were
ligated to lone-linker Lf-SalI, purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S. The products were purified
by phenol/chloroform and Centricon 100. The cDNAs were
digested with SalI and NotI enzymes and cloned into
SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E.
coli host was transformed with the ligation mixture by the
standard chemical method. The average insert size is about
2.3kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 529;
Best Local Similarity 94.4%; Pred. No. 6.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGATCTCTCCAT 18
|||
120 CGACTGATCTCTCCAT 103

RESULT 50
CA533770/c 530 bp mRNA linear EST 18-NOV-2002
LOCUS C0406D02-5N NIA Mouse ES Cell (LIF-) cDNA Library (long) Mus
DEFINITION Musculus cDNA clone NIA: C0406D02 IMAGE:30010213 5', mRNA sequence.
ACCESSION CA533770
VERSION CA533770.1 GI:25065205
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 530)
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K.,
Jaramet, S.A., Boheler, K.R. and Ko, M.S.H.,
Systematic Analyses of NIA Mouse ES Cell (LIF-) cDNA Library (long)
Unpublished (2001)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgun.grc.nia.nih.gov
Plate: C0406 row: D column: 02
Seq primer: M13 Reverse
High quality sequence stop: 530
POLYA-No.

FEATURES

Location/Qualifiers
1..530
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv x 129/Sv-CP"
/db_xref="nia:EST:C0406D02-5N"

ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 530;
Best Local Similarity 94.4%; Pred. No. 6.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGATCTCTCCAT 18
|||
Db 121 CGACTGATCTCTCCAT 104

/db_xref="taxon:10090"
/clone="NIA:C0406D02 IMAGE:30010213"
/tissue_type="ES Cell (LIF-)"
/cell_line="R1 ES cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse ES Cell (LIF-) cDNA Library (long)"
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Dr. Kenneth R. Boheler (National Institute
on Aging, USA). ES cells were cultured without feeder
cells in the absence of LIF for 4 hours or 18 hours.
Equimolar mixture of these RNA samples was used for the
library construction. Double-stranded cDNAs were
synthesized with an Oligo(dT) primer (Invitrogen):
5'-GGACTGATCTAGATCGGAGCGCGCCCTTTT-3' from
21 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker Lf-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.4 kb. The library was constructed
by Yulan Piao (NIA)."

Search completed: January 14, 2005, 17:37:54
Job time : 2450.87 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2005, 10:09:04 ; Search time 301.263 Seconds
(without alignments)
313.645 Million cell updates/sec

Title: US-09-578-453-1

Perfect score: 18
Sequence: 1 CGACTGTGATTCCTCCAT 18

Scoring table: IDENTITY_NUC
Gapc 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database :

N_Geneseq_23Sep04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	2	AAQ08280 Antisense
2	18	100.0	1416	2	AAZ52303 Rat p35 (
3	18	100.0	1627	6	ABK63584 Rat
4	18	100.0	1627	12	ADP72925 Remal tox
5	18	100.0	1786	10	ADB58550 Toxictly-
6	18	100.0	1786	10	ADB51157 Primary r
7	16.4	91.1	93	12	ADL71108 p53 hairp
8	16.4	91.1	390	6	ABK76382 Bacillus
9	16.4	91.1	1173	2	AAO67683 Murine p5
10	16.4	91.1	1173	2	AAZ08528 Murine p5
11	16.4	91.1	1208	12	ABSS1008 Mouse p53
12	16.4	91.1	1208	12	ADJ32551 Mouse p53
13	15.4	85.6	525	2	AAO85528 DNA probe
14	15.4	85.6	584	5	AAAF14040 Aspergill
15	15.4	85.6	608	5	AAAF14040 Aspergill
16	15.4	85.6	795	12	ADN62510 A. thalila
17	15.4	85.6	806	3	AAAC44742 Zea maye
18	15.4	85.6	870	4	AAK92361 Human CDN
19	15.4	85.6	870	4	AAK93899 Human CDN
20	15.4	85.6	870	4	ADL28788 5' end of
21	15.4	85.6	870	12	ADL28788 5' end of

22	15.4	85.6	870	12	ADL30326	Adl30326 3' end of
23	15.4	85.6	1077	3	AAAC4009	Aac4009 Arabidops
24	15.4	85.6	1364	4	AAH14165	Aah14165 Human CDN
25	15.4	85.6	1487	10	ADDB9659	Add6659 Human REM
26	15.4	85.6	1490	4	AAK94874	Aak94874 Human ful
27	15.4	85.6	1490	12	ADL32031	Adl32031 Full leng
28	15.4	85.6	1510	5	AAH64732	Aah64732 Human sec
29	15.4	85.6	1523	5	AAH64771	Aah64771 Human sec
30	15.4	85.6	1525	8	AAV43615	Aav43615 Human sec
31	15.4	85.6	1664	8	ABZ68515	Abz68515 Nucleotid
32	15.4	85.6	1664	12	ADP18659	Adp18659 Human TAT
33	15.4	85.6	1755	2	AAV55741	Aav55741 Human sec
34	15.4	85.6	1755	6	ABQ92067	Abq92067 Human pol
35	15.4	85.6	1892	3	AAAF1834	Aaf1834 Human bre
36	15.4	85.6	2148	4	ABL19347	Ab119347 Drosophi
37	15.4	85.6	2243	4	ABL15659	Ab115659 Drosophi
38	15.4	85.6	3703	4	ABL02040	Ab102040 Drosophi
39	15.4	85.6	4357	4	AAK75168	Aak75168 Human imm
40	15.4	85.6	4532	4	ABL19346	Ab119346 Drosophi
41	15.4	85.6	6446	4	ABL15658	Ab115658 Drosophi
42	15.4	85.6	20000	12	ADQ47193	Ado47193 DNA seque
43	15	83.3	18	10	ACA61261	Aca61261 Rat p35 a
44	15	83.3	1024	3	AAZ97144	Aaz97144 Human pro
45	15	83.3	1065	6	ABN66254	Abn66254 Streptoco
46	15	83.3	11000	6	ABN71527	Abn71527 01
47	14.8	82.2	93	12	ADL71108	Adl71108 p53 hairp
48	14.8	82.2	281	12	ADQ05153	Adq05153 Soybean h
49	14.8	82.2	417	6	ABL66464	Ab166464 Lung canc
50	14.8	82.2	417	6	ABL65765	Ab165765 Lung canc
51	14.8	82.2	417	6	ABN93583	Abn93583 Gene #81
52	14.8	82.2	474	9	ACH32462	Ach32462 Human end
53	14.8	82.2	781	6	AAI45845	Aai45845 Human sec
54	14.8	82.2	791	4	AAI24530	Aai24530 Human bre
55	14.8	82.2	908	12	ADU42388	Adj42388 Plant CDN
56	14.8	82.2	966	4	ABA89005	Ab89005 Escherich
57	14.8	82.2	1002	4	AAAS9713	Aas9713 Propionib
58	14.8	82.2	1002	8	ACF64642	Acf64642 Propionib
59	14.8	82.2	1167	5	AAH46482	Aah46482 Human rib
60	14.8	82.2	1170	4	AAH14899	Aah14899 Human CDN
61	14.8	82.2	1185	4	AAH14287	Aah14287 Human CDN
62	14.8	82.2	1185	12	ADO20164	Ado20164 Human PRO
63	14.8	82.2	1188	6	AAAD24384	Aad24384 Human RNA
64	14.8	82.2	2388	4	ABL23309	Ab123309 Drosophi
65	14.8	82.2	2489	6	ABST78877	Abst78877 E. coli C
66	14.8	82.2	2489	10	ADH80444	Adh80444 Escherich
67	14.8	82.2	2498	4	ABA89004	Ab89004 Escherich
68	14.8	82.2	2541	6	AAI45816	Aai45816 Human sec
69	14.8	82.2	2557	4	ABL29399	Ab129399 Drosophi
70	14.8	82.2	2832	4	ABL05208	Ab105208 Drosophi
71	14.8	82.2	4388	4	ABL23308	Ab123308 Drosophi
72	14.8	82.2	4681	4	ABL23398	Ab123398 Drosophi
73	14.8	82.2	4725	4	ABL29322	Ab129322 Drosophi
74	14.8	82.2	42035	8	AAAS9562	Aas9562 Propionib
75	14.8	82.2	42035	8	AAAS9562	Aas9562 Propionib
76	14.8	82.2	110000	6	ABX08336	Abx08336 11
77	14.8	82.2	110000	6	ABX08336	Abx08336 12
78	14.8	82.2	110000	12	ADU25965	Adu25965 11
79	14.8	82.2	110000	12	ADU25965	Adu25965 12
80	14.8	82.2	110000	12	ADN97989	Adn97989 11
81	14.8	82.2	110000	12	ADN97989	Adn97989 12
82	14.8	82.2	110000	12	ADOS0281	Ado50281 11
83	14.8	82.2	110000	12	ADOS0281	Ado50281 12
84	14.4	80.0	35	12	ADN86991	Adn86991 Mouse p53
85	14.4	80.0	331	4	AAI20107	Aai20107 Probe #10
86	14.4	80.0	331	4	AAI45307	Aai45307 Probe #13
87	14.4	80.0	331	4	AAI45307	Aai45307 Probe #13
88	14.4	80.0	331	4	AAI47251	Aai47251 Human bre
89	14.4	80.0	331	4	AAI47251	Aai47251 Human bre
90	14.4	80.0	331	4	AAK3294	Aak3294 Human bon
91	14.4	80.0	331	4	AAK13555	Aak13555 Human liv
92	14.4	80.0	331	4	AAK13555	Aak13555 Human liv
93	14.4	80.0	331	5	AAI05813	Aai05813 Probe #58
94	14.4	80.0	331	6	ABSI3381	Abi3381 Human gen

95	14.4	80.0	376	6	ABO85387	AbB85387	Arabidops	c 168	14.4	80.0	6911	2	AA734177	AA734177	At34177	Corymefor
96	14.4	80.0	413	9	ACH20958	AbC20958	Human adu	c 169	14.4	80.0	8034	4	ABL03008	ABL03008	AbI03008	Drosophila
97	14.4	80.0	427	8	AXH38233	AbX38233	Bovine ES	c 170	14.4	80.0	8674	8	ABX63814	ABX63814	AbX63814	Human CDN
98	14.4	80.0	433	9	ACH21300	AbC21300	Human adu	c 171	14.4	80.0	8878	2	AA703852	AA703852	AdA03852	Human fib
99	14.4	80.0	444	4	AA110868	AbA110868	Probe #80	c 172	14.4	80.0	8878	6	AAAD45341	AAAD45341	AdA45341	Human ant
100	14.4	80.0	444	4	ABAS2519	AbA52519	Human foe	c 173	14.4	80.0	8878	9	AAAD55117	AAAD55117	AdA55117	Human ant
101	14.4	80.0	444	4	AA132128	AbA132128	Probe #81	c 174	14.4	80.0	8878	11	ADPE5553	ADPE5553	AdP5553	Human FBG
102	14.4	80.0	444	4	ABAA2097	AbA2097	Human bre	c 175	14.4	80.0	8878	11	ADPE5567	ADPE5567	AdP5567	Human fib
103	14.4	80.0	444	4	ABAA22308	AbA22308	Probe #77	c 176	14.4	80.0	11840	4	AAK38820	AAK38820	AdK38820	Human imm
104	14.4	80.0	444	4	AAK26235	AbA26235	Human bon	c 177	14.4	80.0	32195	9	ABK42625	ABK42625	AdB42625	Genomic s
105	14.4	80.0	444	4	AAK00783	AbA00783	Human bra	c 178	14.4	80.0	32195	9	ABD60781	ABD60781	AdB60781	Connectiv
106	14.4	80.0	444	4	ABE25828	AbE25828	Human liv	c 179	14.4	80.0	39796	3	AAAC1681	AAAC1681	AdA1681	Nucleoid
107	14.4	80.0	444	5	AA100792	AbA100792	Probe #78	c 180	14.4	80.0	110000	2	AAZ01425_06	AAZ01425_06	Continuation (7 of	
108	14.4	80.0	445	6	ABSO0823	AbS00823	Human gen	c 181	14.4	80.0	349980	5	AAH68531	AAH68531	AdH68531	C glutami
109	14.4	80.0	445	3	AACT71409	AbC71409	Single nu	c 182	14.4	77.8	385	4	AAK55788	AAK55788	AdK55788	Human imm
110	14.4	80.0	455	3	AACT71418	AbC71418	Single nu	c 183	14.4	77.8	387	10	ACD97529	ACD97529	AdC97529	Human col
111	14.4	80.0	455	3	AACT71443	AbC71443	Single nu	c 184	14.4	77.8	455	8	ABZ55728	ABZ55728	AbZ55728	Aspergill
112	14.4	80.0	477	9	ACH21912	AbC21912	Human adu	c 185	14.4	77.8	529	10	ADK57503	ADK57503	AdK57503	Plant DNA
113	14.4	80.0	495	6	ABSG6915	AbG6915	Novel mur	c 186	14.4	77.8	750	4	AAI24883	AAI24883	AdI24883	Human bre
114	14.4	80.0	573	4	AAI17550	AbI17550	Probe #74	c 187	14.4	77.8	852	3	AAAC43068	AAAC43068	AdA43068	Arabidops
115	14.4	80.0	573	4	ABA62483	AbA62483	Human foe	c 188	14.4	77.8	1136	3	AAAC38191	AAAC38191	AdA38191	Arabidops
116	14.4	80.0	573	4	AAI42463	AbI42463	Probe #11	c 189	14.4	77.8	1138	3	AAAC45620	AAAC45620	AdA45620	Arabidops
117	14.4	80.0	573	4	AAK29814	AbA29814	Probe #82	c 190	14.4	77.8	1234	6	ABSG6542	ABSG6542	AdB6542	Mouse Mga
118	14.4	80.0	573	4	AAK36698	AbA36698	Human bon	c 191	14.4	77.8	1395	2	AAAI5662	AAAI5662	AdA15662	Protein p
119	14.4	80.0	573	4	AAK10835	AbA10835	Human bra	c 192	14.4	77.8	1395	10	ADK64853	ADK64853	AdK64853	Disease t
120	14.4	80.0	573	4	ABSG36357	AbS36357	Human liv	c 193	14.4	77.8	2088	12	ADN74420	ADN74420	AdN74420	Thale cre
121	14.4	80.0	573	6	ABSI0699	AbS10699	Human gen	c 194	14.4	77.8	2125	12	ADG20490	ADG20490	AdG20490	Mouse pro
122	14.4	80.0	598	12	ACH73732	AbC73732	Human gen	c 195	14.4	77.8	2430	3	AACT9712	AACT9712	AdC9712	Arabidops
123	14.4	80.0	864	6	ABZ14669	AbZ14669	Thalecres	c 196	14.4	77.8	3109	12	ADJ87641	ADJ87641	AdJ87641	Nervous s
124	14.4	80.0	864	12	ADDO1806	AdO1806	Thalecres	c 197	14.4	77.8	4914	8	ABV76029	ABV76029	AbV76029	Mouse pla
125	14.4	80.0	920	2	ABN98651	AbN98651	Arabidops	c 198	14.4	77.8	13813	4	AAK89790	AAK89790	AdK89790	Human dig
126	14.4	80.0	1093	2	AA795761	AbT95761	Arabidops	c 199	14.4	77.8	13904	10	ADCG87172	ADCG87172	AdC87172	Human GPC
127	14.4	80.0	1093	3	AAH67333	AbH67333	C glucam	c 200	14.4	77.8	42048	2	AAK71918	AAK71918	AdK71918	Human imm
128	14.4	80.0	1098	5	AAH67333	AbH67333	C glucam	c 201	14.4	77.8	43226	2	AAK60263	AAK60263	AdK60263	Nucleic a
129	14.4	80.0	1118	3	AAH67333	AbH67333	C glucam	c 202	13.8	76.7	24	12	ADH23526	ADH23526	AdH23526	PCR prime
130	14.4	80.0	1464	3	AAAC7238	AbA7238	Arabidops	c 203	13.8	76.7	25	9	ADCK15038	ADCK15038	AdC15038	Human mic
131	14.4	80.0	1629	12	ADBT76867	AbD76867	Human CDN	c 204	13.8	76.7	34	9	ADA02250	ADA02250	AdA02250	Mouse car
132	14.4	80.0	1850	2	AAV12203	AbV12203	Cyclochrom	c 205	13.8	76.7	34	10	ADH31989	ADH31989	AdH31989	Mouse car
133	14.4	80.0	1850	2	AAV12203	AbV12203	Cyclochrom	c 206	13.8	76.7	49	12	ADH32608	ADH32608	AdH32608	Mouse car
134	14.4	80.0	1850	2	AAH22442	AbH22442	Zebrafish	c 207	13.8	76.7	50	12	ADM32609	ADM32609	AdM32609	HEIV2 Gre
135	14.4	80.0	1850	6	ABQ74193	AbQ74193	Zebrafish	c 208	13.8	76.7	51	4	AAH90180	AAH90180	AdH90180	Human cto
136	14.4	80.0	1850	6	ABD24484	AbD24484	Zebrafish	c 209	13.8	76.7	60	6	ABN44820	ABN44820	AdB44820	Human spl
137	14.4	80.0	1862	4	AAK51626	AbK51626	Human pol	c 210	13.8	76.7	103	3	ADB76015	ADB76015	AdB76015	Tomato pl
138	14.4	80.0	1883	11	ADPE5676	AdP5676	Human fib	c 211	13.8	76.7	200	4	ABA70112	ABA70112	AdA70112	Human foe
139	14.4	80.0	1883	11	ADPE5552	AdP5552	Human fib	c 212	13.8	76.7	200	4	AAI50231	AAI50231	AdI50231	Probe #18
140	14.4	80.0	1885	4	AAAS40978	AbA40978	Human fib	c 213	13.8	76.7	200	4	ABA56898	ABA56898	AdA56898	Probe #15
141	14.4	80.0	1918	11	ADPE5102	AdP5102	Human fib	c 214	13.8	76.7	200	4	AAK44226	AAK44226	AdK44226	Human bon
142	14.4	80.0	1918	11	ADPE5031	AdP5031	Human fib	c 215	13.8	76.7	200	4	AAK18327	AAK18327	AdK18327	Human bra
143	14.4	80.0	1935	4	ABL27239	AbL27239	Drosophila	c 216	13.8	76.7	200	4	ABSA3883	ABSA3883	AdB3883	Human liv
144	14.4	80.0	2111	4	AAAS41533	AbA41533	CDNA enco	c 217	13.8	76.7	201	6	ABSI8462	ABSI8462	AdB8462	Human gen
145	14.4	80.0	2120	5	AAK51624	AbK51624	Human pol	c 218	13.8	76.7	201	2	AAH85659	AAH85659	AdH85659	Human bin
146	14.4	80.0	2120	5	AAH78178	AbH78178	DNA enco	c 219	13.8	76.7	202	2	ABV03890	ABV03890	AdV03890	Human pro
147	14.4	80.0	2484	5	AAH78178	AbH78178	DNA enco	c 220	13.8	76.7	230	2	AAQ77548	AAQ77548	AdQ77548	Human gen
148	14.4	80.0	2553	4	AAK52608	AbK52608	Human pol	c 221	13.8	76.7	246	6	ABK80560	ABK80560	AdB80560	Bacillus
149	14.4	80.0	2553	4	AAK52609	AbK52609	Human pol	c 222	13.8	76.7	300	6	ABL75469	ABL75469	AdL75469	Corn tass
150	14.4	80.0	2553	4	AAK52610	AbK52610	Human pol	c 223	13.8	76.7	311	2	AAV89340	AAV89340	AdV89340	EST clone
151	14.4	80.0	2586	5	AAAS67724	AbA67724	DNA enco	c 224	13.8	76.7	320	9	ACCT2381	ACCT2381	AdC2381	Human CDN
152	14.4	80.0	2664	12	ADMT2852	AdM2852	Thale cre	c 225	13.8	76.7	337	6	ABN21293	ABN21293	AdB21293	Human ORF
153	14.4	80.0	2774	10	ADDS5861	AdD5861	Thalecres	c 226	13.8	76.7	343	2	AAO61319	AAO61319	AdO61319	Human bra
154	14.4	80.0	2822	12	ADQ24895	AdQ24895	Human sof	c 227	13.8	76.7	402	6	AB199683	AB199683	AdB199683	Mouse isc
155	14.4	80.0	2856	4	AAH14376	AbH14376	Human CDN	c 228	13.8	76.7	405	5	ABV43041	ABV43041	AdV43041	Human pro
156	14.4	80.0	2910	4	ABL03009	AbL03009	Drosophila	c 229	13.8	76.7	405	5	ABV43180	ABV43180	AdV43180	Human pro
157	14.4	80.0	3032	4	ABL07891	AbL07891	Drosophila	c 230	13.8	76.7	419	5	ABV07910	ABV07910	AdV07910	Human pro
158	14.4	80.0	3152	11	ADM39520	AdM39520	Human CDN	c 231	13.8	76.7	431	4	AAK64059	AAK64059	AdK64059	Human imm
159	14.4	80.0	3397	10	ADDA6454	AdD6454	Rat gene	c 232	13.8	76.7	432	3	AAAC32187	AAAC32187	AdA32187	Human sec
160	14.4	80.0	3417	10	ADBE2654	AdB2654	Human CDN	c 233	13.8	76.7	440	9	ACH48224	ACH48224	AdC48224	Human lun
161	14.4	80.0	4043	6	ABX91986	AbX91986	Lung spec	c 234	13.8	76.7	442	9	ACH50637	ACH50637	AdC50637	Human mam
162	14.4	80.0	4374	6	ABL27238	AbL27238	Drosophila	c 235	13.8	76.7	443	5	ABV13059	ABV13059	AdV13059	Human pro
163	14.4	80.0	4392	10	ADDE60010	AdD60010	Human gen	c 236	13.8	76.7	452	6	AB181752	AB181752	AdB181752	Human ova
164	14.4	80.0	4392	10	ADDE60010	AdD60010	Human gen	c 237	13.8	76.7	459	8	ACA35048	ACA35048	AdA35048	Prokaryot
165	14.4	80.0	4967	5	AAH60957	AbH60957	Human can	c 238	13.8	76.7	460	2	AA731598	AA731598	AdA731598	Alpha cha
166	14.4	80.0	4967	5	ADL63650	AdL63650	Human ova	c 239	13.8	76.7	464	8	ABX36089	ABX36089	AdB36089	Bovine ES
167	14.4	80.0	5299	4	ABL07890	AbL07890	Drosophila	c 240	13.8	76.7	474	4	ABA57487	ABA57487	AdA57487	Human foe

C 241	13.8	76.7	474	4	ABA57842	AbA57842 Human foe	C 314	13.8	76.7	1041	3	AAZ99484	AAZ99484 cDNA enco
C 242	13.8	76.7	474	4	AAI37038	AAI37038 Probe #57	C 315	13.8	76.7	1041	6	AAAD4074	AAAD4074 Pumpkin 2
C 243	13.8	76.7	474	4	AAI37433	AAI37433 Probe #61	C 316	13.8	76.7	1046	3	AAC99067	AAC99067 Human pan
C 244	13.8	76.7	474	4	ABA27182	ABA27182 Probe #56	C 317	13.8	76.7	1053	6	ABN98389	ABN98389 Arabidops
C 245	13.8	76.7	474	4	AAK31131	AAK31131 Human bon	C 318	13.8	76.7	1086	3	AAC38741	AAC38741 Arabidops
C 246	13.8	76.7	474	4	AAK31554	AAK31554 Human bon	C 319	13.8	76.7	1146	6	AAZ23611	AAZ23611 Canine p5
C 247	13.8	76.7	474	4	AAK05916	AAK05916 Human bra	C 320	13.8	76.7	1166	2	AAZ42037	AAZ42037 Human end
C 248	13.8	76.7	474	4	AAK05529	AAK05529 Human bra	C 321	13.8	76.7	1174	6	AAZ23609	AAZ23609 Canine p5
C 249	13.8	76.7	474	4	ABK30811	ABK30811 Human liv	C 322	13.8	76.7	1207	3	AAC45575	AAC45575 Arabidops
C 250	13.8	76.7	474	4	ABK31238	ABK31238 Human liv	C 323	13.8	76.7	1210	3	AAC33305	AAC33305 Arabidops
C 251	13.8	76.7	474	6	ABK05883	ABK05883 Human gen	C 324	13.8	76.7	1230	8	ACA53938	ACA53938 Prokaryot
C 252	13.8	76.7	474	6	ABK06308	ABK06308 Human gen	C 325	13.8	76.7	1272	8	ACA47824	ACA47824 Prokaryot
C 253	13.8	76.7	475	5	ABV37833	ABV37833 Human pro	C 326	13.8	76.7	1296	3	ACA51827	ACA51827 Arabidops
C 254	13.8	76.7	486	4	AAK07135	AAK07135 DNA enco	C 327	13.8	76.7	1305	8	ACA32204	ACA32204 Prokaryot
C 255	13.8	76.7	487	4	AAK41066	AAK41066 Zea may	C 328	13.8	76.7	1332	3	AAC42359	AAC42359 Arabidops
C 256	13.8	76.7	491	3	ACH33647	ACH33647 Human end	C 329	13.8	76.7	1373	12	ADH70091	ADH70091 Human Vbe
C 257	13.8	76.7	494	10	ADF85661	ADF85661 Human ade	C 330	13.8	76.7	1395	6	ABZ12919	ABZ12919 Arabidops
C 258	13.8	76.7	501	4	ABA60737	ABA60737 Human ade	C 331	13.8	76.7	1395	12	ADN73174	ADN73174 Ttale cre
C 259	13.8	76.7	501	4	AAI40628	AAI40628 Probe #93	C 332	13.8	76.7	1431	8	ACC51103	ACC51103 Human Dop
C 260	13.8	76.7	501	4	AAK34912	AAK34912 Human bon	C 333	13.8	76.7	1431	8	ACC51104	ACC51104 Human Dop
C 261	13.8	76.7	501	4	AAK09021	AAK09021 Human bra	C 334	13.8	76.7	1457	10	ADC32127	ADC32127 Human nov
C 262	13.8	76.7	501	4	ABK34668	ABK34668 Human liv	C 335	13.8	76.7	1463	3	AAC52279	AAC52279 Arabidops
C 263	13.8	76.7	514	4	AAFI1929	AAFI1929 Human bre	C 336	13.8	76.7	1493	3	AAC77153	AAC77153 Human ORF
C 264	13.8	76.7	514	4	AAK47359	AAK47359 Human bre	C 337	13.8	76.7	1506	6	ABA91373	ABA91373 Human bre
C 265	13.8	76.7	514	6	ABK63960	ABK63960 Human bre	C 338	13.8	76.7	1513	6	ABK46131	ABK46131 cDNA enco
C 266	13.8	76.7	514	10	ABT33172	ABT33172 Human tum	C 339	13.8	76.7	1513	8	ACC50153	ACC50153 Breast ca
C 267	13.8	76.7	514	11	ADL33079	ADL33079 Human bre	C 340	13.8	76.7	1513	10	ADB75292	ADB75292 Prostate
C 268	13.8	76.7	514	12	ADK44369	ADK44369 Human cDN	C 341	13.8	76.7	1518	3	AAZ49454	AAZ49454 Mouse Vol
C 269	13.8	76.7	516	11	ACH96986	ACH96986 Klebsell	C 342	13.8	76.7	1545	12	ADP04788	ADP04788 Sea equir
C 270	13.8	76.7	519	4	AAFI1770	AAFI1770 Human bre	C 343	13.8	76.7	1548	3	AAAC48768	AAAC48768 Arabidops
C 271	13.8	76.7	519	6	AAK47200	AAK47200 Human bre	C 344	13.8	76.7	1551	3	AAC32972	AAC32972 Arabidops
C 272	13.8	76.7	519	4	ABT08855	ABT08855 Human bre	C 345	13.8	76.7	1552	5	AAZ50890	AAZ50890 Human rec
C 273	13.8	76.7	519	6	ABK63801	ABK63801 Human bre	C 346	13.8	76.7	1626	5	AAK66192	AAK66192 DNA enco
C 274	13.8	76.7	519	10	ABT33013	ABT33013 Human tum	C 347	13.8	76.7	1677	5	AAK77114	AAK77114 DNA enco
C 275	13.8	76.7	519	11	ADL29290	ADL29290 Human bre	C 348	13.8	76.7	1677	8	ABK63140	ABK63140 Human cDN
C 276	13.8	76.7	519	12	ADK44210	ADK44210 Human cDN	C 349	13.8	76.7	1715	10	ADFI17114	ADFI17114 Lettuce .c
C 277	13.8	76.7	522	5	AAK76168	AAK76168 DNA enco	C 350	13.8	76.7	1722	2	AAT00402	AAT00402 Squale
C 278	13.8	76.7	529	5	ABV15936	ABV15936 Human pro	C 351	13.8	76.7	1724	12	ADH13717	ADH13717 Human ENZ
C 279	13.8	76.7	544	10	ACD93052	ACD93052 Human col	C 352	13.8	76.7	1734	2	AAQ97800	AAQ97800 Clone pCB
C 280	13.8	76.7	565	5	ABV45736	ABV45736 Human pro	C 353	13.8	76.7	1734	6	ABT99373	ABT99373 Mouse 18c
C 281	13.8	76.7	589	5	ABV56962	ABV56962 Human pro	C 354	13.8	76.7	1741	3	AAC32892	AAC32892 Arabidops
C 282	13.8	76.7	598	3	AAK32814	AAK32814 Arabidops	C 355	13.8	76.7	1745	3	AAK45529	AAK45529 Arabidops
C 283	13.8	76.7	620	10	ADD33713	ADD33713 Mouse mit	C 356	13.8	76.7	1794	10	ADC03223	ADC03223 Human nov
C 284	13.8	76.7	640	10	ADG37560	ADG37560 Aspergill	C 357	13.8	76.7	1831	4	ABT04375	ABT04375 Drosophi
C 285	13.8	76.7	649	4	AAH72227	AAH72227 Human cer	C 358	13.8	76.7	1841	12	ADH13220	ADH13220 Human mal
C 286	13.8	76.7	687	5	ABV28360	ABV28360 Human pro	C 359	13.8	76.7	1853	2	AAV73011	AAV73011 Human adu
C 287	13.8	76.7	687	5	ABV22545	ABV22545 Human pro	C 360	13.8	76.7	1853	6	ABQ092036	ABQ092036 Human pol
C 288	13.8	76.7	690	10	ADD33716	ADD33716 Mouse mit	C 361	13.8	76.7	1893	4	AAH16343	AAH16343 Human cDN
C 289	13.8	76.7	700	6	ABK46132	ABK46132 cDNA enco	C 362	13.8	76.7	1938	2	AAZ42120	AAZ42120 Human end
C 290	13.8	76.7	758	4	AAH05160	AAH05160 Human cDN	C 363	13.8	76.7	1956	2	AAQ88712	AAQ88712 Aspergill
C 291	13.8	76.7	765	3	AAZ47241	AAZ47241 HTLV-1 Ta	C 364	13.8	76.7	2000	10	ADC08567	ADC08567 Rice DNA
C 292	13.8	76.7	765	3	AAZ47276	AAZ47276 HLA-A2/HT	C 365	13.8	76.7	2002	3	AACT7836	AACT7836 Human can
C 293	13.8	76.7	765	3	AAZ56677	AAZ56677 TCR alpha	C 366	13.8	76.7	2027	2	AAK006649	AAK006649 Human sec
C 294	13.8	76.7	765	3	AAZ56667	AAZ56667 TCR alpha	C 367	13.8	76.7	2067	8	ACA339054	ACA339054 Prokaryot
C 295	13.8	76.7	778	4	AAH71164	AAH71164 Human cer	C 368	13.8	76.7	2070	12	ADL04080	ADL04080 DNA enco
C 296	13.8	76.7	792	6	ABZ14145	ABZ14145 Arabidops	C 369	13.8	76.7	2137	4	ABL15573	ABL15573 Drosophi
C 297	13.8	76.7	796	2	AAI29757	AAI29757 D10 singl	C 370	13.8	76.7	2163	6	ABN66445	ABN66445 Streptoco
C 298	13.8	76.7	859	4	AAI35629	AAI35629 Human mus	C 371	13.8	76.7	2179	6	ABQ99371	ABQ99371 Human cod
C 299	13.8	76.7	859	4	ABK58617	ABK58617 cDNA enco	C 372	13.8	76.7	2199	10	ADBS88258	ADBS88258 Toxicity-
C 300	13.8	76.7	859	12	ADU28344	ADU28344 Human mus	C 373	13.8	76.7	2199	10	ADBS85175	ADBS85175 Primary r
C 301	13.8	76.7	870	5	AAK79520	AAK79520 DNA enco	C 374	13.8	76.7	2199	10	ADH85175	ADH85175 Rat equal
C 302	13.8	76.7	870	5	AAK72842	AAK72842 DNA enco	C 375	13.8	76.7	2199	12	ADP72506	ADP72506 Renal tox
C 303	13.8	76.7	905	6	ABK65888	ABK65888 Lung canc	C 376	13.8	76.7	2225	4	ABL03473	ABL03473 Drosophi
C 304	13.8	76.7	905	6	ABN93926	ABN93926 Gene #424	C 377	13.8	76.7	2238	12	ADU39529	ADU39529 Plant cDN
C 305	13.8	76.7	936	5	AAK50137	AAK50137 DNA enco	C 378	13.8	76.7	2307	8	ABX71087	ABX71087 Novel hum
C 306	13.8	76.7	949	3	AAK31537	AAK31537 Human ade	C 379	13.8	76.7	2420	4	AAH72859	AAH72859 Human cer
C 307	13.8	76.7	949	3	AAK21259	AAK21259 Human low	C 380	13.8	76.7	2420	5	ABV25077	ABV25077 Human pro
C 308	13.8	76.7	949	10	ABZ96953	ABZ96953 Human nuc	C 381	13.8	76.7	2465	4	AAH16721	AAH16721 Human cDN
C 309	13.8	76.7	949	11	ABD20802	ABD20802 Human pul	C 382	13.8	76.7	2510	11	ADM02669	ADM02669 Human cDN
C 310	13.8	76.7	964	3	AAK02697	AAK02697 Human col	C 383	13.8	76.7	2513	3	AAH14259	AAH14259 Human cDN
C 311	13.8	76.7	967	4	AAH33694	AAH33694 Human col	C 384	13.8	76.7	2568	8	ABT18130	ABT18130 Aspergill
C 312	13.8	76.7	1016	6	ABK34682	ABK34682 Human cDN	C 385	13.8	76.7	2683	8	ABT19944	ABT19944 Aspergill
C 313	13.8	76.7	1032	6	ABK67322	ABK67322 Breast sp	C 386	13.8	76.7	2718	4	AAI37575	AAI37575 Human mus

387	13.8	76.7	2718	8	ABX60563	Abx60563 cDNA enco
388	13.8	76.7	2718	12	ADJ31313	Adj31313 Human mus
389	13.8	76.7	2792	12	ABL05249	AbL05249 Drosophi
390	13.8	76.7	2825	6	AAD29162	Aad29162 Sheep K2.
391	13.8	76.7	2825	8	ACC49767	Acc49767 Ovine K2.
392	13.8	76.7	2831	4	ABL06391	AbL06391 Drosophi
393	13.8	76.7	2859	12	ADP99210	Adp99210 Human tra
394	13.8	76.7	3019	5	AAH83874	Aah83874 DNA enco
395	13.8	76.7	3104	4	AAH15749	Aah15749 Human cdn
396	13.8	76.7	3114	4	AAH15789	Aah15789 Human cdn
397	13.8	76.7	3146	1	AAH81713	Aah81713 Clone con
398	13.8	76.7	3166	11	ADMO2054	Admo2054 Human cdn
399	13.8	76.7	3194	12	ADQ24917	Adq24917 Human sof
400	13.8	76.7	3267	12	ADN72212	Adn72212 Thale cre
401	13.8	76.7	3349	8	ABZ73720	Abz73720 Secreted
402	13.8	76.7	3349	10	ADCC0594	Adcc0594 Human sec
403	13.8	76.7	3349	10	ABT16868	Abt16868 Human sec
404	13.8	76.7	3349	10	ABZ67316	Abz67316 Human sec
405	13.8	76.7	3350	8	ABZ73719	Abz73719 Secreted
406	13.8	76.7	3350	10	ADCC0593	Adcc0593 Human sec
407	13.8	76.7	3350	10	ABT16867	Abt16867 Human sec
408	13.8	76.7	3350	10	ABZ67315	Abz67315 Human sec
409	13.8	76.7	3443	12	ADN04827	Adn04827 Act1pso
410	13.8	76.7	3734	3	AAA35138	Aaa35138 Human ade
411	13.8	76.7	3734	6	AAAF21260	Aaf21260 Human low
412	13.8	76.7	3734	6	ABK83777	Abk83777 Human cdn
413	13.8	76.7	3734	11	ABD20803	Abd20803 Human nuc
414	13.8	76.7	3734	11	ADP17112	Adp17112 Letuce c
415	13.8	76.7	3839	10	AAH72798	Aah72798 Human cer
416	13.8	76.7	3856	5	ADL65359	Adl65359 Human ova
417	13.8	76.7	3900	4	ABL04374	AbL04374 Drosophi
418	13.8	76.7	4345	4	AAK75170	Aak75170 Human imm
419	13.8	76.7	4357	4	AAK75169	Aak75169 Human imm
420	13.8	76.7	4377	2	AAV23493	Aav23493 Pseudom
421	13.8	76.7	4377	3	AAAI3904	Aai3904 Pseudom
422	13.8	76.7	4377	3	AAAF30869	Aaf30869 Pseudom
423	13.8	76.7	4377	6	AAD22881	Aad22881 P. alca
424	13.8	76.7	4377	6	ABT43709	Abt43709 Molecu
425	13.8	76.7	4518	4	ABLO3472	AbL03472 Drosophi
426	13.8	76.7	4539	4	ABLO3472	AbL03472 Drosophi
427	13.8	76.7	4667	3	AAH35139	Aah35139 Human ade
428	13.8	76.7	4683	3	AAAF21261	Aaf21261 Human low
429	13.8	76.7	4683	10	ABZ69555	Abz69555 Human nuc
430	13.8	76.7	4683	10	ABD20804	Abd20804 Human pul
431	13.8	76.7	4683	11	ABD20804	Abd20804 Human pul
432	13.8	76.7	4831	4	ABL06390	AbL06390 Drosophi
433	13.8	76.7	4920	4	ABL15572	AbL15572 Drosophi
434	13.8	76.7	4928	4	ABL06823	AbL06823 Drosophi
435	13.8	76.7	5240	2	AAV48268	Aav48268 P. chryo
436	13.8	76.7	5377	4	ABL16036	AbL16036 Drosophi
437	13.8	76.7	5439	4	ABLO5248	AbL05248 Drosophi
438	13.8	76.7	5526	5	ABAI9882	Abai9882 Human ner
439	13.8	76.7	5825	5	ACH03918	Ach03918 Human cdn
440	13.8	76.7	6102	10	ADBS5284	Adbs5284 Rat gene
441	13.8	76.7	6102	10	ADBS5300	Adbs5300 Rat gene
442	13.8	76.7	6102	10	ADBS5292	Adbs5292 Rat gene
443	13.8	76.7	6102	10	ADBS5304	Adbs5304 Rat gene
444	13.8	76.7	6102	10	ADBS5286	Adbs5286 Rat gene
445	13.8	76.7	6102	10	ADBS5288	Adbs5288 Rat gene
446	13.8	76.7	6393	2	AAZ23079	Aaz23079 Human plu
447	13.8	76.7	6394	8	ACC50261	Acc50261 Breast ca
448	13.8	76.7	6438	8	ABZ73721	Abz73721 Secreted
449	13.8	76.7	6438	10	ADCC0595	Adcc0595 Human sec
450	13.8	76.7	6438	10	ABT16869	Abt16869 Human sec
451	13.8	76.7	6438	10	ABZ67317	Abz67317 Human sec
452	13.8	76.7	6483	4	AAH57440	Aah57440 Human int
453	13.8	76.7	6513	12	ADOI9559	Adoi9559 Human pro
454	13.8	76.7	6641	12	ADQ24481	Adq24481 Human sof
455	13.8	76.7	6930	6	ABO93533	Abg93533 Human dif
456	13.8	76.7	7063	4	AAH24651	Aah24651 Nucleotid
457	13.8	76.7	7450	12	ADQ19096	Adq19096 Human sof
458	13.8	76.7	7465	4	ABLO5016	AbL05016 Drosophi
459	13.8	76.7	7473	10	ADP81503	Adp81503 Leukaemia

460	13.8	76.7	7473	12	ADQ23420	Adq23420 Human sof
461	13.8	76.7	9596	4	AAK66614	Aak66614 Human imm
462	13.8	76.7	9596	4	AAK66613	Aak66613 Human imm
463	13.8	76.7	9903	4	AAI05500	Aai05500 Human rep
464	13.8	76.7	10785	4	ABA07367	Abao7367 Human pan
465	13.8	76.7	10785	4	AAK91098	Aak91098 Human pan
466	13.8	76.7	10785	4	AAK54411	Aak54411 Human dig
467	13.8	76.7	10785	5	AAK30060	Aak30060 Human lun
468	13.8	76.7	10785	10	ADB33397	Adb33397 Human nov
469	13.8	76.7	11740	11	ADN97366	Adn97366 B. jichen
470	13.8	76.7	13315	4	ABA07371	Abao7371 Human pan
471	13.8	76.7	13315	4	AAK91102	Aak91102 Human dig
472	13.8	76.7	13315	4	AAK54415	Aak54415 Human dig
473	13.8	76.7	13315	5	AAK30064	Aak30064 Human lun
474	13.8	76.7	13315	10	ADB33401	Adb33401 Human nov
475	13.8	76.7	14141	4	ABLO6822	AbL06822 Drosophi
476	13.8	76.7	15654	12	ADM32597	Adm32597 Human par
477	13.8	76.7	15654	12	ADM32595	Adm32595 Human par
478	13.8	76.7	15654	12	ADM32596	Adm32596 Human par
479	13.8	76.7	16084	4	AAK69442	Aak69442 Human imm
480	13.8	76.7	16084	4	AAK82881	Aak82881 Human imm
481	13.8	76.7	19437	4	AAK85311	Aak85311 Human imm
482	13.8	76.7	20261	4	ABA07368	Abao7368 Human pan
483	13.8	76.7	20261	4	AAK91099	Aak91099 Human dig
484	13.8	76.7	20261	5	AAK54412	Aak54412 Human imm
485	13.8	76.7	20261	5	AAK30061	Aak30061 Human lun
486	13.8	76.7	20261	10	ADP33398	Adp33398 Human nov
487	13.8	76.7	24187	4	AAK59529	Aak59529 Propionib
488	13.8	76.7	24187	8	ACF64458	Acf64458 Propionib
489	13.8	76.7	24550	4	AAK59608	Aak59608 Propionib
490	13.8	76.7	24550	8	ACF64537	Acf64537 Propionib
491	13.8	76.7	32223	5	ABAI7368	Abai7368 Human ner
492	13.8	76.7	32223	5	ABAI7368	Abai7368 Human ner
493	13.8	76.7	32835	9	ADAO2948	Ada02948 Human nek
494	13.8	76.7	32835	10	ADB72686	Adb72686 Human nek
495	13.8	76.7	32835	10	ADCB85428	Adcb85428 Human nek
496	13.8	76.7	32835	12	ADM74543	Adm74543 Human car
497	13.8	76.7	51001	12	ADU10262	AdU10262 Human ger
498	13.8	76.7	55986	9	ADA02741	Ada02741 Mouse Trf
499	13.8	76.7	55986	10	ADB72479	Adb72479 Mouse Trf
500	13.8	76.7	55986	10	ADC85221	Adc85221 Mouse Trf

ALIGNMENTS

RESULT 1	
ID	AAQ88280 standard; cDNA; 18 BP.
XX	AAQ88280;
AC	
XX	
DT	25-MAR-2003 (revised)
DT	06-DEC-1995 (first entry)
XX	
DE	Antisense oligonucleotide to inhibit p53 gene expression.
XX	
KW	Antisense; recombinant virus; vector; p53; binding site; Huntington's;
KW	neuronal degeneration; Parkinson's; Alzheimer's;
KW	amyotrophic lateral sclerosis; ss.
XX	
OS	Synthetic.
XX	
PN	MO950916-A1.
XX	
PD	13-APR-1995.
XX	
PF	29-SEP-1994; 94WO-FR001142.
XX	
PR	04-OCT-1993; 93FR-00011774.
XX	
PA	(RHON) RHONE POULENC RORER SA.
XX	

PI Mallet J, Revah F, Stutzmann J;
XX WPI; 1995-155257/20.
XX
PT Treating and preventing neurodegenerative diseases - with cpds. that
PT inhibit p53 activity, partic. an anti-sense sequence or viral vector.
XX
PS Example 1; Page 10; 21pp; French.
XX
CC The sequence of an antisense oligonucleotide used either solely or as
CC part of a recombinant viral vector to inhibit the expression of the p53
CC gene. The oligonucleotide may be used in conjunction, either as naked DNA
CC or on the same recombinant viral vector, with the p53 binding sequence
CC (AAQ88279). The viral vector may also encode a mutated form of p53 which
CC antagonises p53 activity. These oligonucleotides and recombinant viruses
CC can be used in compositions to treat and prevent neuronal degeneration,
CC Huntington's, Parkinson's or Alzheimer's diseases or amyotrophic lateral
CC sclerosis. (updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 18 BP; 4 A; 6 C; 3 G; 5 T; 0 U; 0 Other;
Query Match 100.0%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CGACTGTGATCTCTCCAT 18
Db 1 CGACTGTGATCTCTCCAT 18
RESULT 2
AAZ52303/c
ID AAZ52303 standard; cDNA; 1416 BP.
XX
AC AAZ52303;
XX
DT 24-JUL-2000 (first entry)
XX
DE Rat p35 (p53 isoform) cDNA.
XX
KM Rat p53 isoform; p35; reverse transcription-PCR; marker; hypoxia;
KM myocardial infarction; cell proliferation; cytostatic; proliferative; ss.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 24..838
FT /tag= a
FT /product= "Rat p35 protein"
FT /transl_except= (pos: 759..760, aa: Arg)
FT /transl_except= (pos: 821..823, aa: Ala)
XX
PN MO200022127-A1.
XX
PD 20-APR-2000.
XX
PE 07-OCT-1999; 99WO-US023319.
XX
PR 09-OCT-1998; 98US-0103849P.
XX
PA (BGM) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Dell'acqua G, Mann MJ, Dzaou VJ;
XX
DR WPI; 2000-317984/27.
DR P-PSDB; AAV70713.
XX
PT Novel isoform of p53 useful as a marker of myocardial infarction and for
PT controlling cellular proliferation.
XX
PS Claim 7; Fig 2; 25pp; English.
XX
CC The present cDNA sequence encodes the rat p53 isoform, p35 derived from

CC rat cardiac myocytes by reverse transcription-PCR of RNA. This p53
CC isoform is truncated to eliminate a substantial portion of the C-terminal
CC end of the protein. p35 is useful as a marker of myocardial infarction
CC (by the indication of hypoxia) and in the control of cell proliferation
CC both in vivo and in vitro. It was able to increase the transactivation
CC achieved through co-transfection with wild type p53. The combination of
CC p35 with wild type p53 produce an enhanced inhibition of primary
CC incorporation at 24 hours compared to wild type p53 alone
XX
SQ Sequence 1416 BP; 312 A; 432 C; 323 G; 349 T; 0 U; 0 Other;
Query Match 100.0%; Score 18; DB 3; Length 1416;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CGACTGTGATCTCTCCAT 18
Db 41 CGACTGTGATCTCTCCAT 24
RESULT 3
ABK63584/c
ID ABK63584 standard; cDNA; 1627 BP.
XX
AC ABK63584;
XX
DT 18-JUN-2002 (first entry)
XX
DE Rat sequence differentially expressed in response to a hepatotoxin #1491.
XX
KM Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KM differential expression; centrilobular necrosis; steatosis.
XX
OS Rattus norvegicus.
XX
PN MO200210453-A2.
XX
PD 07-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-US023872.
XX
PR 31-JUL-2000; 2000US-0222040P.
PR 02-NOV-2000; 2000US-0244880P.
PR 11-MAY-2001; 2001US-0230029P.
PR 15-MAY-2001; 2001US-0230645P.
PR 22-MAY-2001; 2001US-0232336P.
PR 06-JUN-2001; 2001US-0235798P.
PR 13-JUN-2001; 2001US-0237457P.
PR 19-JUN-2001; 2001US-0238884P.
PR 09-JUL-2001; 2001US-0303459P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX
DR WPI; 2002-241625/29.
XX
PT Predicting toxic effects of compounds or the progression of these toxic
PT effects by determining the changes in gene expression in tissues or cells
PT exposed to the toxin and comparing these to gene expression in unexposed
PT tissues or cells.
XX
PS Claim 1; SEQ ID NO 1491; 239pp; English.
XX
CC The invention relates to methods for predicting toxic effects of
CC compounds or the progression of these toxic effects by determining the
CC global changes in gene expression in tissues or cells exposed to the
CC toxin and comparing these to gene expression in unexposed tissues or
CC cells. Also included are methods of predicting at least one toxic effect
CC of a compound or progression of a toxic effect, preferably the
CC hepatotoxicity of a compound, comprising detecting the level of
CC expression in a tissue or cell sample exposed to the compound of two or
CC more genes listed in the specification, where differential expression of

CC the genes is indicative of at least one toxic effect or progression. The
CC method can also be used to identify an agent which modulates the toxic
CC response and predict cellular pathways that a compound modulates in a
CC cell. The methods utilize a set of at least two probes (on a solid
CC support in kit form), where each of the probes comprises a sequence that
CC specifically hybridises to a gene listed in the specification, a computer
CC system comprising a database containing information identifying the
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC set of genes comprising at least two genes listed in the specification,
CC and a user interface to view the information used to present information
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cell exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell
CC sample that has been exposed to a compound or agent. Hepatotoxicity is
CC characterised by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent

XX SQ Sequence 1627 BP; 374 A; 486 C; 380 G; 387 T; 0 U; 0 Other;
Query Match 100.0%; Score 18; DB 6; Length 1627;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCAT 18
Db 41 CGACTGTGAATCCTCCAT 24

RESULT 4

ADP72925/C.
ID ADP72925 standard; DNA; 1627 BP.

XX AC ADP72925;

XX DT 26-AUG-2004 (first entry)

XX DB Renal toxin progression gene marker #1514.

XX ds; toxic effect; gene expression profile; kidney tissue;
XX differential gene expression; toxicity progression; toxicity marker;
XX drug screening; toxicity assay; kidney pathology; nephritis;
XX kidney necrosis; glomerular injury; tubular injury;
XX local segmental glomerulosclerosis.

XX OS Rattus norvegicus.

XX PN WO2004048598-A2.

XX PD 10-JUN-2004.

XX PF 24-NOV-2003; 2003WO-US037556.

XX PR 22-NOV-2002; 2002US-00301856.

XX (GENE-) GENE LOGIC INC.

XX PA Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
XX PI Elashoff M;

XX WPI; 2004-460771/43.

XX Predicting (the progression of) a toxic effect of a compound, for
XX monitoring the progression of renal disease states, comprises preparing a
XX gene expression profile of a kidney tissue or cell sample exposed to the
XX compound.

XX Claim 11; SEQ ID NO 1514; 266bp; English.

CC The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The
CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC kidney necrosis, glomerular and tubular injury, or focal segmental
CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC predicting or modelling the toxic response of a test compound, for
CC monitoring the progression of renal disease states, for identifying genes
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at http://wipo.int/pub/published_pct_sequences).

XX SQ Sequence 1627 BP; 374 A; 486 C; 380 G; 387 T; 0 U; 0 Other;
Query Match 100.0%; Score 18; DB 12; Length 1627;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCAT 18
Db 41 CGACTGTGAATCCTCCAT 24

RESULT 5

ADB58550/C
ID ADB58550 standard; DNA; 1786 BP.

XX AC ADB58550;

XX DT 04-DEC-2003 (first entry)

XX DE Toxicity-related gene, SEQ ID 3576.

XX KM Toxic; toxin; gene expression profile; hepatotoxicity; liver;
XX drug screening; toxicity assay; ds.

XX OS Unidentified.

XX PN WO2003064624-A2.

XX PD 07-AUG-2003.

XX PF 31-JAN-2003; 2003WO-US003194.

XX PR 31-JAN-2002; 2002US-00060087.

XX PR 15-MAR-2002; 2002US-0364045P.

XX PR 15-MAR-2002; 2002US-0364055P.

XX PR 30-DEC-2002; 2002US-0436643P.

XX (GENE-) GENE LOGIC INC.

XX PA Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX PI WPI; 2003-689530/65.

XX Predicting a toxic effect of a compound, useful in identifying toxicity
XX markers in liver tissues or cells for drug screening and toxicity assays,
XX comprises preparing gene expression profile of tissue or cells exposed to
XX the compound.

XX Claim 1; SEQ ID NO 3576; 1156bp; English.

XX

PT cells into the subject.

PS Example 1; Page 37; 93pp; English.

CC The invention relates to a novel method for introducing into a subject a
CC population of stem cells having partial or complete loss of function of a
CC target gene by introducing a nucleic acid construct encoding a short
CC hairpin RNA (shRNA) into stem cells to generate transfected stem cells
CC and introducing the transfected stem cells into the subject. The
CC transfected stem cells propagate within the subject and retain partial to
CC complete loss of function of the target gene. The invention has anti-HIV
CC and cytostatic activity, and may have a use in cell therapy. The
CC composition may be administered intravenously or intraarterially, or into
CC the liver (via portal vein injection), the peritoneal cavity, the kidney
CC capsule or the bone marrow. The composition and methods are useful for
CC performing RNA interference in stem cells and for treating diseases (e.g.
CC HIV or cancer) associated with the expression of a target gene in a
CC population of cells. The present sequence represents a shRNA nucleic acid
CC of the invention.

SQ Sequence 93 BP; 26 A; 27 C; 19 G; 21 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 12; Length 93;

Best Local Similarity 94.4%; Pred. No. 75;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18
Db 66 CGACTGTGAATCCTCCAT 49

RESULT 8

ABK76382/c
ID ABK76382 standard; DNA; 390 BP.

XX ABK76382;

DT 13-AUG-2002 (first entry)

DE Bacillus licheniformis genomic sequence tag (GST) #3673.

XX Differential gene expression; genomic sequenced tag; GST;

KW altered culture condition; environmental stress;

OS physiological provocation; ds.

XX Bacillus licheniformis.

OS WO200229113-A2.

XX 11-APR-2002.

PF 05-OCT-2001; 2001WO-US031437.

XX 06-OCT-2000; 2000US-00680598.

PR 27-MAR-2001; 2001US-0279526P.

XX (NOVO) NOVOZYMES BIOTECH INC.

PA (NOVO) NOVOZYMES AS.

XX Berka R, Clausen IG;

PI WPI; 2002-416684/44.

XX Monitoring differential expression of several genes in first Bacillus

PT cell relative to expression of same genes in one or more second Bacillus

PT cells, by using substrate containing Bacillus genomic sequenced tag

PT array.

XX Claim 4; SEQ ID NO 3673; 200pp; English.

XX The invention describes a method of monitoring differential expression of

CC genes in a first Bacillus cell relative to expression of the genes in

CC other Bacillus cells, comprising hybridising labelled nucleic acid probes

CC isolated from Bacillus cells to a substrate containing array of Bacillus

CC genomic sequenced tags (GST), examining the array, and determining

CC relative gene expression by an observed hybridisation reporter signal of

CC a spot in the array. The method is useful for measuring the expression of

CC genes in a first Bacillus cell relative to expression of the same genes

CC in one or more second Bacillus cells. The method is useful for monitoring

CC global expression of several genes from a Bacillus cell, discovering new

CC genes, identifying possible functions of unknown open reading frames and

CC monitoring gene copy number variation and stability. Monitoring changes

CC in expression of genes may be used to provide a representation of the way

CC in which Bacillus cells adapt to changes in culture conditions,

CC environmental stress or other physiological provocation. Extensive follow

CC up characterisation is unnecessary, when one spot on an array equals one

CC gene or one open reading frame, since sequence information is available.

CC This sequence represents a genomic sequence tag (GST) used in the method

CC of the invention. Note: The sequence data for this patent did not form

CC part of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 390 BP; 96 A; 98 C; 96 G; 100 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 6; Length 390;

Best Local Similarity 94.4%; Pred. No. 85;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGTGAATCCTCCAT 18
Db 120 CGACTGTGAATCCTCCAT 103

RESULT 9

AAQ67883/c
ID AAQ67883 standard; DNA; 1173 BP.

XX AAQ67883;

DT 25-MAR-2003 (revised)

DT 23-MAR-1995 (first entry)

XX Murine p53 DNA.

XX Polymerase chain reaction; primer; amplify; NVVAC; ALVAC; recombinant;

KW murine; interleukin-2; IL-2; PRW825; pmu-1; PBS-SK; pmu151; TK vector;

KW plasmid; vaccinia; H6 promoter; amplify; primer; antigenic response;

KW polymerase chain reaction; poxvirus; psd542; immunological response;

XX pathogen; human; interferon; IFN; ss.

XX Synthetic.

OS WO9416716-A1.

XX 04-AUG-1994.

XX 21-JAN-1994; 94WO-US000888.

XX 21-JAN-1993; 93US-00007115.

XX 19-JAN-1994; 94US-00184009.

XX (VIR-) VIROGENETICS CORP.

XX Paoletti E, Tartaglia J, Cox WI;

XX WPI; 1994-263767/32.

XX Attenuated recombinant virus used for cancer therapy - comprises DNA

PT encoding cytokine and/or tumour associated antigen.

XX Example 31; Fig 38; 232pp; English.

XX This sequence represents the wildtype murine p53 gene from the

CC translation initiation codon to the stop codon. This sequence was used in

CC the construction of an ALVAC-based recombinant virus containing the wild-

CC type murine p53 gene. Three PCR fragments were generated. Fragment I was

CC produced using primers MM080 and MM081 and plasmid PRW825 as a template.

CC The 228 bp amplified fragment contains the H6 promoter and the 5'-most

PT Novel complex of protein-protein interactions in adipocyte cells for
PT identifying compounds that modulate the protein-protein interactions and
PT useful for treating obesity and metabolic disorders.

PS Claim 2, Page, 125pp; English.

XX The invention relates to a complex of protein-protein interactions
CC (forming a protein-protein interaction map, PIM) in adipocyte cells as
CC defined in the specification, or polynucleotides in adipocytes encoding
CC for the polypeptides. Also included are a recombinant cell expressing the
CC interacting polypeptides and a method of selecting a modulating compound
CC in adipocyte cells, by cultivating a recombinant host cell on a selective
CC medium containing a modulating compound and a reporter gene the
CC expression of which is toxic for the recombinant host cell which is
CC transformed with two vectors, where the first vector comprises a
CC polynucleotide encoding a first hybrid polypeptide and DNA binding domain
CC and the second vector comprising a polynucleotide encoding a second
CC hybrid polypeptide and an activating domain that activates the toxic
CC reporter gene, when the first and second hybrid polypeptides interact and
CC selecting the modulating compound which inhibits the growth of the
CC recombinant host cell (i.e. using the yeast two-hybrid system). The
CC complexes are useful for identifying compounds that modulate the protein-
CC protein interactions and useful for treating obesity and metabolic
CC disorders e.g. non-insulin dependent diabetes mellitus, NIDDM. The
CC compound isolated by the method is useful for treating and preventing
CC obesity or metabolic diseases. The interactions between the proteins of
CC the complex further define a set of selected interacting domains, SID.
CC The present sequence encodes a member of the protein complex of the
CC invention, used as the bait protein in the yeast two- hybrid assay. Note:
CC The present sequence was not displayed in the specification but was
CC obtained from its Genbank entry by the indexer

XX Sequence 1208 BP; 283 A; 362 C; 314 G; 249 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 6; Length 1208;
Best Local Similarity 94.4%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGGAATCCTCCAT 18
Db 52 CGACTGGAATCCTCCAT 35

RESULT 12
ADJ32551/c
ID ADJ32551 standard; DNA; 1208 BP.

AC ADJ32551;

DT 22-APR-2004 (first entry)

XX Mouse p53 DNA.

XX Integrin; cell proliferation; cancer; melanoma; adenoma; lymphoma;
KW myeloma; carcinoma; glioma; plasmocytoma; sarcoma; thymoma; leukaemia;
KW brain tumour; cytosarcoma; mouse; gene; ds.

XX Mus musculus.

XX Key Location/Qualifiers
FH CDS 26..1198
FT /tag= a
FT /product= "Mouse p53 protein"

XX US2003224993-A1.

XX 04-DEC-2003.

XX 17-MAR-2003; 2003US-00392113.

XX 12-OCT-2000; 2000US-0239705P.

XX 24-OCT-2000; 2000US-0242812P.
PR 12-OCT-2001; 2001WO-US032127.

PR 15-MAR-2002; 2002US-0365078P.

XX (LAND/) LAND H.
PA (DELE/) DELEU L.

XX Land H, Deleu L;

DR WPI, 2004-154528/15.
P-P8DB; ADJ32552.

XX Reducing the proliferation of a cancer cell involves inhibiting ligand
PT binding to an integrin receptor on the cancer cell, where the integrin
PT receptor comprises an integrin.

PS Disclosure; SEQ ID NO 32; 161pp; English.

XX The present invention relates to compositions and methods for reducing
CC the proliferation of cancer cells through interaction with integrins. The
CC invention is useful for reducing the proliferation of cancer cells e.g.
CC melanoma, adenoma, lymphoma, myeloma, carcinoma, glioma, plasmocytoma,
CC sarcoma, thymoma, leukaemia, skin cancer, retinal cancer, breast cancer,
CC prostate cancer, colon cancer, esophageal cancer, stomach cancer,
CC pancreas cancer, brain tumours, lung cancer, ovarian cancer, cervical
CC cancer, hepatic cancer, gastrointestinal cancer, and head and neck cancer
CC cells. The invention is also useful for identifying a therapeutic target
CC which involves assaying potential reagent for activity. The present
CC sequence is mouse p53 DNA.

XX Sequence 1208 BP; 283 A; 362 C; 314 G; 249 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 12; Length 1208;
Best Local Similarity 94.4%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGGAATCCTCCAT 18
Db 52 CGACTGGAATCCTCCAT 35

RESULT 13
AAQ85528/c
ID AAQ85528 standard; DNA; 396 BP.

AC AAQ85528;

DT 10-JAN-1996 (first entry)

XX DNA probe 34 detects DNA-protein complex in immortal cells.

XX DNA-protein complex; detection; proliferation; tumour formation;
KW diagnose; malignancy; biopsy; probe; ss.

XX Synthetic.

XX WO9502701-A1.

XX 26-JAN-1995.

XX 13-JUL-1994; 94WO-EP002307.

XX 15-JUL-1993; 93DE-04323727.

XX (BOE) BOEHRINGER MANNHEIM GMBH.

XX Abken HU, Albert W, Jungfer H;

XX WPI; 1995-067344/09.

XX New DNA-protein complex characteristic of cells with unlimited
PT proliferation capacity - and its components and derived antibodies,
PT useful in diagnosing malignant tumours.

XX Claim 1, Page 75; 106pp; German.

XX The DNA shown is found in human or animal cells that have an unlimited
CC capacity for unlimited cell proliferation or tumor formation. They have
CC no ability to promote immortalisation of the cells, and are usually found
CC in a DNA-protein complex in the cell cytoplasm. The DNA is useful for
CC detection of these complexes, and diagnosis of malignant tumours.
CC Differentiation between cells with unlimited and only transient
CC proliferation is possible. (See also AA085493-54)

XX
SQ Sequence 396 BP; 105 A; 50 C; 138 G; 99 T; 0 U; 4 Other;
Query Match 85.6%; Score 15.4; DB 2; Length 396;
Best Local Similarity 94.1%; Pred. No. 3e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
|||
Db 302 GACTGTGAATCCTCCAT 286

RESULT 14
AAH05726
ID AAH05726 standard; cDNA; 525 BP.
XX
AC AAH05726;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:2561.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PE 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
PI Oka T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 1; SEQ ID NO 2561; 2537bp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification; where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence; where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in

CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 525 BP; 84 A; 166 C; 151 G; 121 T; 0 U; 3 Other;
Query Match 85.6%; Score 15.4; DB 4; Length 525;
Best Local Similarity 94.1%; Pred. No. 3.1e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCA 17
|||
Db 434 CGACTGTGAATCCTCCA 450

RESULT 15
AAFI4040
ID AAFI4040 standard; cDNA; 584 BP.
XX
XX AAFI4040;
XX
DT 13-MAR-2001 (first entry)
XX
DE Aspergillus oryzae EST SEQ ID NO:6563.
XX
XX Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus oryzae.
XX
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PE 22-MAR-2000; 2000WO-US007781.
XX
PR 22-MAR-1999; 99US-00273623.
XX
PR (NOVO) NOVO NORDISK BIOTECH INC.
XX (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
PI WPI; 2000-594572/56.
XX
XX Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags.
XX
XX Claim 88; Page 2686; 3161pp; English.
XX
XX The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring the
CC global expression of genes from FF cells allows the production potential
CC of the microorganisms to be improved. New genes may be discovered,
CC possible functions of unknown open reading frames can be identified and
CC gene copy number variation and stability can be monitored. The expression
CC of genes can be used to study how FF cells adapt to changes in culture

AAc44742
ID AAc44742 standard; DNA; 806 BP.
XX
AC AAc44742;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 43964.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic; pathway;
KW promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
PD
PD 06-SEP-2000.
XX
PP 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128645P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0134449P.
PR 30-APR-1999; 99US-0135048P.
PR 30-APR-1999; 99US-0135407P.
PR 04-MAY-1999; 99US-0134848P.
PR 05-MAY-1999; 99US-0134858P.
PR 06-MAY-1999; 99US-0134866P.
PR 07-MAY-1999; 99US-0134877P.
PR 11-MAY-1999; 99US-0134863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138647P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140695P.
PR 29-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144848P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145066P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145126P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149226P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149729P.

```
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157535P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161456P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

```
Query March 85.6%; Score 15.4; DB 3; Length 806;
Best Local Similarity 94.1%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 GACTGTGAATCCTCCAT 18
Db 446 GACTGTGAATCCTCCAT 462
```

```
RESULT 19
AAK92361
ID AAK92361 standard; cDNA; 870 BP.
XX
AC AAK92361;
XX
DT 06-NOV-2001 (first entry)
XX
```

```
DE Human cDNA 5'-end sequence, SEQ ID NO: 821.
XX
XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX Homo sapiens.
XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
XX Claim 2; SEQ ID NO 821, 1380pp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is the nucleotide sequence of the 5'-end of
CC a cDNA provided in the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in CD-
CC ROM format directly from EPO
XX
XX Sequence 870 BP; 161 A; 264 C; 231 G; 213 T; 0 U; 1 Other;
SQ
```

```
Query March 85.6%; Score 15.4; DB 4; Length 870;
Best Local Similarity 94.1%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CGACTGTGAATCCTCCA 17
Db 417 CGACTGTGAATCCTCCA 433
```

```
RESULT 20
AAK93899
ID AAK93899 standard; cDNA; 870 BP.
```

```
XX
AC AAK93899;
XX
DT 06-NOV-2001 (first entry)
XX
XX Human cDNA clone representative sequence, SEQ ID NO: 2359.
XX
XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX Homo sapiens.
XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
```

PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 XX
 PR 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX
 PS Example 11; SEQ ID NO 2359; 1380bp + Sequence Listing; English.
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesized by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence was used as the representative sequence
 CC from a human clone which was used in homology searches to identify the
 CC clone. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in CD-ROM format directly from
 CC BPO
 CC
 SQ Sequence 870 BP; 161 A; 264 C; 231 G; 213 T; 0 U; 1 Other;
 XX
 XX
 Query Match 85.6%; Score 15.4; DB 4; Length 870;
 Best Local Similarity 94.1%; Pred. No. 3.2e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CGACTGTGATTCCTCCA 17
 Db 417 CGACTGTGATTCCTCCA 433
 XX
 XX
 RESULT 21
 ADL28788
 ID ADL28788 standard; cDNA; 870 BP.
 XX
 AC ADL28788;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE 5' end of a human cDNA molecule SegID 821.
 XX
 KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1396543-A2.
 XX
 PD 10-MAR-2004.
 XX
 PF 07-JUL-2000; 2003EP-00025638.
 XX
 PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183865.
 PR 07-JUL-2000; 2000EP-00114089.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2004-204755/20.
 XX

PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 XX
 XX
 PS Disclosure; SEQ ID NO 821; 1340bp; English.
 CC This invention relates to a novel primers useful for synthesizing full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction.
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polynucleotide sequence is the
 CC 5' end of a full length human cDNA sequence of the invention.
 XX
 SQ Sequence 870 BP; 161 A; 264 C; 231 G; 213 T; 0 U; 1 Other;
 XX
 XX
 Query Match 85.6%; Score 15.4; DB 12; Length 870;
 Best Local Similarity 94.1%; Pred. No. 3.2e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CGACTGTGATTCCTCCA 17
 Db 417 CGACTGTGATTCCTCCA 433
 XX
 XX
 RESULT 22
 ADL30326
 ID ADL30326 standard; cDNA; 870 BP.
 XX
 AC ADL30326;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE 3' end of a representative human cDNA cluster SegID 2359.
 XX
 KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1396543-A2.
 XX
 PD 10-MAR-2004.
 XX
 PF 07-JUL-2000; 2003EP-00025638.
 XX
 PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183865.
 PR 07-JUL-2000; 2000EP-00114089.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2004-204755/20.
 XX
 PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 XX
 XX
 PS Example 18; SEQ ID NO 2359; 1340bp; English.
 CC This invention relates to a novel primers useful for synthesizing full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction.
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polynucleotide sequence is the
 CC 3' end of a representative human DNA cluster of the invention.

XX SQ Sequence 870 BP; 161 A; 264 C; 231 G; 213 T; 0 U; 1 Other;
Query Match 85.6%; Score 15.4; DB 12; Length 870;
Best Local Similarity 94.1%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 CGACTGTGATCTCTCCA 17
417 CGACTGTGATCTCTCCA 433
Db
RESULT 23
AAC40009/C
ID AAC40009 standard; DNA; 1077 BP.
XX AC AAC40009;
XX 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26716.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-012180P.
PR 09-MAR-1999; 99US-012548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135623P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.


```

PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149358P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155466P.
PR 24-SEP-1999; 99US-0155658P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158028P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159328P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 16-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

```

```

Query Match 85.6%; Score 15.4; DB 3; Length 1077;
Best Local Similarity 94.1%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 2 GACTGTGATCTCTCAT 18
Db 825 GACTGTGATCTCTCAT 809

RESULT 24
AAH14165
ID AAH14165 standard; cDNA, 1364 BP.
XX
AC AAH14165;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SFG ID NO:11397.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUN-2000; 2000BP-00116126.
XX
PR 29-JUN-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 11397; 2537bp + Sequence listing; English.
XX
XX
CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 1364 BP; 267 A; 415 C; 370 G; 312 T; 0 U; 0 Other;

```

```

Query Match 85.6%; Score 15.4; DB 4; Length 1364;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 CGACTGTGAATCCTCCA 17
 DB 434 CGACTGTGAATCCTCCA 450

RESULT 25

ADD69659 standard; cDNA, 1487 BP.

ADD69659;

15-JAN-2004 (first entry)

Human REMAP CDNA - SEQ ID 88.

human; receptor and membrane-associated protein; REMAP; cytosolic; antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective; cerebroprotective; anti-HIV; antiallergic; antiinflammatory; thymometric; cell proliferative; cancer; atherosclerosis; neurological; epilepsy; Huntington's disease; stroke; immune; inflammatory; AIDS; allergy; developmental; hypothyroidism; Cushing's syndrome; infection; ss; gene.

Homo sapiens.

WO2003048305-A2.

12-JUN-2003.

13-NOV-2002; 2002WO-US036759.

15-NOV-2001; 2001US-0333097P.

14-DEC-2001; 2001US-0340542P.

18-DEC-2001; 2001US-0342166P.

11-JAN-2002; 2002US-0347580P.

14-JAN-2002; 2002US-0348687P.

(INCY-) INCYTE GENOMICS INC.

Duggan BM, Yang J, Gietzen KJ, Lee SY, Tang YT, Azimzal Y, Chawla NK, Warren BA, Barroso I, Becha SD, Yue H, Lehr-Mason PM, Thangavelu K, Lee S, Emerling BM, Kable AB, Khare R, Baughn MR, Gandhi AR, Tran UK, Richardson TW, Margis JP, Lal PG, Forsythe LJ, Lee EA, Swarnakar A, Kallick DA, Griffin JA, Elliott VS, Gervay AE, Hatalia AA, Ison CH, Jin P, Jiang X, Jackson AA, Bhatia U, Buttrill JD, Blake UT, Ho A, Zheng W, Gao J, WPI; 2003-513744/48. P-PSDB; ADD69612.

New human receptors and membrane-associated proteins (REMAP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant REMAP expression e.g. cancer, AIDS, atherosclerosis, or infections.

Claim 5; SEQ ID NO 88; 298bp; English.

The invention relates to a novel isolated polypeptide comprising a human receptor and membrane-associated protein (REMAP). The polypeptide of the invention demonstrates cytosolic, antiarteriosclerotic, anticonvulsant, neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic, antiinflammatory and thymometric activities and may be useful for treating and diagnosing various disorders including those which are cell proliferative such as cancer and atherosclerosis, neurological including epilepsy, Huntington's disease and stroke, immune/inflammatory particularly AIDS and allergies and developmental such as hypothyroidism and Cushing's syndrome, as well as infections. The current sequence is that of the human REMAP CDNA of the invention.

Sequence 1487 BP; 319 A; 442 C; 392 G; 334 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 1487;
 Best Local Similarity 94.1%; Pred. No. 3.3e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCA 17
 DB 542 CGACTGTGAATCCTCCA 558

RESULT 26

AAK94874 standard; cDNA, 1490 BP.

AAK94874;

06-NOV-2001 (first entry)

Human full-length cDNA, SEQ ID NO: 4064.

Human, full length cDNA; cDNA synthesis; oligo-capping; ss.

Homo sapiens.

EP130094-A2.

05-SEP-2001.

07-JUL-2000; 2000EP-00114089.

08-JUL-1999; 99JP-00194486.

11-JAN-2000; 2000JP-00118774.

02-MAY-2000; 2000JP-00183765.

(HELI-) HELIX RES INST.

Oka T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y, Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H, WPI; 2001-524255/58. P-PSDB; AAM93912.

830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation.

Claim 8; SEQ ID NO 4064; 1380bp + Sequence listing; English.

The invention relates to primers for synthesizing full length cDNA clones. 830 CDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5' and 3' ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesized by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a full length human cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO

Sequence 1490 BP; 288 A; 455 C; 399 G; 348 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 4; Length 1490;
 Best Local Similarity 94.1%; Pred. No. 3.3e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGAATCCTCCA 17
 DB 417 CGACTGTGAATCCTCCA 433

RESULT 27

ADL32031 standard; cDNA, 1490 BP.

```

XX AC ADL32031;
XX XX 20-MAY-2004 (first entry)
XX XX Full length human cDNA clone Segid 4064.
XX DE human; medicine; signal transduction; glycoprotein; transcription;
XX KW oligo-capping method; ss; gene.
XX XX Homo sapiens.
XX OS BP1396543-A2.
XX PN 10-MAR-2004.
XX PD 07-JUL-2000; 2003EP-00025638.
XX XX 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183865.
XX PR 07-JUL-2000; 2000EP-00114089.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,
XX DR WPI; 2004-204755/20.
XX DR P-PSDB; ADL32032.
XX PT New oligonucleotide primers (830 CDNAs) useful for synthesizing full
XX PT length human cDNAs.
XX PS Example 1; SEQ ID NO 4064; 1340bp; English.
XX XX This invention relates to a novel primers useful for synthesizing full
XX CC length cDNA molecules that encode human proteins. Specifically, it refers
XX CC to secretory or membrane proteins that are potential therapeutic agents/
XX CC target molecules in the field of medicine, and in particular genes
XX CC encoding proteins that are associated with signal transduction,
XX CC glycoproteins and transcription. The present invention describes a method
XX CC for efficiently cloning a full length human cDNA from both the 5' and 3'
XX CC ends using the oligo-capping method. This polynucleotide sequence is a
XX CC full length human cDNA clone of the invention.
XX SQ Sequence 1490 BP; 288 A; 455 C; 399 G; 348 T; 0 U; 0 Other;
XX
XX Query Match 85.6%; Score 15.4; DB 12; Length 1490;
XX Best Local Similarity 94.1%; Pred. No. 3.3e+02;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGAATCTCTCA 17
Db 417 CGACTGTGATTCCTCCA 433

```

RESULT 28
AAH64732
ID AAH64732 standard; cDNA; 1510 BP.
XX
AC AAH64732;
XX
XX 11-SEP-2001 (first entry)
XX DE Human secreted protein cDNA, SEQ ID NO: 8.
XX KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
XX KW GENSET; ss.
XX OS Homo sapiens.
XX PN WO200142451-A2.

```

XX PD 14-JUN-2001.
XX XX 07-DEC-2000; 2000WO-IB001938.
XX PF 08-DEC-1999; 99US-0169629P.
XX PR 06-MAR-2000; 2000US-0187470P.
XX XX (GENSET) GENSET.
XX PA Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX PI WPI; 2001-367870/38.
XX DR P-PSDB; AAG89129.
XX DR Full length GENSET human nucleic acids encoding potentially secreted
XX PT proteins, useful in gene therapy and vaccination against a variety of
XX PT diseases, and for diagnosis of those diseases.
XX PS Claim 7; Page 566-568; 921pp; English.
XX XX The invention relates to full length GENSET human nucleic acids encoding
XX CC potentially secreted proteins. The nucleic acids and the polypeptides
XX CC they encode may be used in the prevention, treatment and diagnosis of
XX CC diseases associated with inappropriate GENSET gene expression. For
XX CC example, they be used to treat disorders associated with decreased GENSET
XX CC gene expression by rectifying mutations or deletions in a patient's
XX CC genome that affect the activity of GENSET or by supplementing the
XX CC patient's own production of GENSET polypeptides. Conversely, antisense
XX CC nucleic acid molecules may be administered to down regulate GENSET
XX CC expression by binding with the cells' own genes and preventing their
XX CC expression. The sense and antisense nucleic acids may also be used as DNA
XX CC probes in diagnostic assays to detect and quantitate the presence of
XX CC similar nucleic acid sequences in samples, and hence to determine which
XX CC patients may be in need of restorative therapy. The GENSET polypeptides
XX CC may be used as antigens in the production of antibodies and in assays to
XX CC identify modulators (agonists and antagonists) of GENSET polypeptide
XX CC expression and activity. The present sequence is a GENSET nucleic acid of
XX CC the invention.
XX SQ Sequence 1510 BP; 307 A; 452 C; 404 G; 347 T; 0 U; 0 Other;
XX
XX Query Match 85.6%; Score 15.4; DB 5; Length 1510;
XX Best Local Similarity 94.1%; Pred. No. 3.3e+02;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGAATCTCTCA 17
Db 405 CGACTGTGATTCCTCCA 421

```

RESULT 29
AAH64771
ID AAH64771 standard; cDNA; 1523 BP.
XX
AC AAH64771;
XX
XX 11-SEP-2001 (first entry)
XX DE Human secreted protein cDNA, SEQ ID NO: 47.
XX KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
XX KW GENSET; ss.
XX OS Homo sapiens.
XX PN WO200142451-A2.
XX PD 14-JUN-2001.
XX PF 07-DEC-2000; 2000WO-IB001938.
XX PR 08-DEC-1999; 99US-0169629P.

```

PR 06-MAR-2000; 2000US-0187470P.
XX
XX (GENSET ) GENSET.
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX
DR WPI; 2001-367870/38.
XX
DR P-PSDB; AAG89168.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
XX
PS Claim 7; Page 612-614; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased GENSET
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patient's own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention
XX
SQ Sequence 1523 BP; 308 A; 462 C; 404 G; 349 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 5; Length 1523;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATCCTCCA 17
Db 432 CGACTGTGATCCTCCA 448
RESULT 30
AAV43615
ID AAV43615 standard; DNA; 1535 BP.
XX
AC AAV43615;
XX
DT 24-SEP-1998 (first entry)
XX
DE Human secreted protein 15 encoding DNA.
XX
XX Secreted protein; human; cell proliferation; cytokine activity;
XX tissue growth; cellular differentiation; regeneration; activin; inhibin;
XX chemotactic; haemostatic; thrombolytic; tumour inhibition;
XX anti-inflammatory activity; biomarker; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 220..1224
FT /*tag= a
FT /product= "human secreted protein"
XX
XX WO9825959-A2.
XX
XX 18-JUN-1998.
XX
XX 11-DEC-1997; 97WO-US022787.
XX

```

```

PR 11-DEC-1996; 96US-0032757P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX
PI Escobedo J, Hu Q, Garcia P, Williams LT, Kothakota S;
XX
DR WPI; 1998-348453/30.
XX
DR P-PSDB; AAM63695.
XX
PT Secreted human polypeptides - having cytokine, cell proliferation or
PT differentiation, activin or inhibin, tumour inhibition or anti-
PT inflammatory activities.
XX
XX
PS Claim 6; Page 42; 78pp; English.
XX
CC This DNA encodes a human secreted protein. The specification provides
CC secreted protein sequences (AAM63681 to AAM63699) encoded by the nucleic
CC acid sequences shown in AAV43601 to AAV43619. The invention provides a
CC method of identifying a secreted polypeptide which is modified by rough
CC microsome. The secreted proteins can be used in assays to determine
CC biological activities, such as cytokine, cell proliferation, or cellular
CC differentiation activities, tissue growth or regeneration, activin or
CC inhibin activity, chemotactic or chemokinetic activity, haemostatic or
CC thrombolytic activity, receptor/ligand activity, tumour inhibition, or
CC anti-inflammatory activity. The proteins can also be used as biomarkers,
CC to identify tissues or cell types which express the proteins, or a stage-
CC or disease-specific alteration in protein expression. They can be used in
CC protein interaction assays, to identify ligands or binding proteins.
CC Compounds which affect the biological activities of the secreted proteins
CC or their ability to interact with specific ligands can be identified
CC using the proteins in screening assays. The proteins and antibodies that
CC bind specifically to the protein can also be used to design diagnostic
CC tests and therapeutic compositions for diseases which may be associated
CC with altered expression of these proteins. Fusion proteins comprising,
CC e.g. signal sequences or transmembrane domains of the proteins can be
CC used to target other protein domains to cellular membrane or they can be
CC secreted extracellularly
XX
SQ Sequence 1535 BP; 313 A; 463 C; 408 G; 351 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 2; Length 1535;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATCCTCCA 17
Db 435 CGACTGTGATCCTCCA 451
RESULT 31
ABZ68515
ID ABZ68515 standard; DNA; 1664 BP.
XX
AC ABZ68515;
XX
DT 22-APR-2003 (first entry)
XX
DE Nucleotide sequence of human CGLI polypeptide.
XX
XX Human; congenital generalized lipodystrophy protein; CGLI; 11q13;
XX chromosome 11; congenital generalized lipodystrophy; lipodystrophy;
XX diabetes; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 345..1541
FT /*tag= a
FT /product= "GLI"
XX
XX FR824332-A1.
XX
XX 08-NOV-2002.
XX

```

```

XX 04-MAY-2001; 2001FR-00006037.
XX
XX 04-MAY-2001; 2001FR-00006037.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (NAGE-) CENT NAT GENOTYPAGE.
XX
PI Magre J, Capeau J, Lathrop M, Delapine M;
XX
XX WPI; 2003-142459/14.
XX P-PSDB; ABP97511.
XX
XX Nucleic acid encoding a congenital generalized lipodystrophy gene cg11
XX PT and mutations of that gene, useful to prevent and treat congenital
XX PT generalized lipodystrophy and obesity.
XX
XX Claim 4; Page 11; 115pp; French.
XX
XX The present sequence encodes a human congenital generalized lipodystrophy
XX CC protein, designated CG1. The CG1 gene is localised at 11q13 on
XX CC chromosome 11. CG1 is responsible for congenital generalized
XX CC lipodystrophy. CG1 polypeptides and polynucleotides are used for
XX CC preventing or treating lipodystrophy or diabetes. CG1 polypeptides are
XX CC also useful as immunogens for raising antibodies.
XX
XX Sequence 1664 BP; 382 A; 481 C; 428 G; 373 T; 0 U; 0 Other;
SQ
Query Match 85.6%; Score 15.4; DB 8; Length 1664;
Best Local Similarity 94.1%; Pred. No. 3.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATCTCTCCA 17
Db 560 CGACTGTGATCTCTCCA 576
RESULT 32
ADP18659
ID ADP18659 standard; cDNA; 1664 BP.
XX
XX ADP18659;
XX
XX 26-AUG-2004 (first entry)
DT
XX Human TAT417 cDNA used to treat cancer SeqID 19.
DE
XX tumour; growth inhibitory; cytotoxic; cytostatic; gene therapy; cancer;
XX KW human; TAT; gene; ss.
XX
XX Homo sapiens.
OS
XX WO2004045516-A2.
XX
XX 03-JUN-2004.
XX
XX 13-NOV-2003; 2003WO-US036298.
XX
XX 15-NOV-2002; 2002US-0426847P.
XX PR 06-DEC-2002; 2002US-0431250P.
XX PR 31-DEC-2002; 2002US-0437344P.
XX
XX (GERTH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Frantz G, Goddard A, Gonzalez J, Gurney AL;
XX PI Polakis P, Polson A, Wood WI, Wu TD, Zhang Z;
XX
XX WPI; 2004-420515/39.
XX P-PSDB; ADP18680.
XX
XX New antibody that binds to tumor-associated antigenic target polypeptide
XX PT (TAT), useful in preparing a composition for diagnosing or treating
XX PT tumor.

```

```

XX Claim 1; SEQ ID NO 19; 183pp; English.
XX
XX This invention relates to novel isolated antibodies that are useful for
XX CC the diagnosis and treatment of a tumour. Specifically, it refers to DNA
XX CC that is detectably labelled and conjugated to a growth inhibitory or
XX CC cytotoxic agent and comprises toxins, antibiotics, radioactive isotopes
XX CC or nucleolytic enzymes. The present invention refers to the toxin
XX CC maytansinoid or calicheamicin, where the host cell is one of chinese
XX CC hamster ovary (CHO), yeast or Escherichia coli cell. Furthermore, these
XX CC pharmaceutical compositions are cytostatic and can be used for gene
XX CC therapy purposes to treat various cancers. This polynucleotide sequence
XX CC is a human TAT cDNA sequence of the invention.
XX
XX Sequence 1664 BP; 382 A; 481 C; 428 G; 373 T; 0 U; 0 Other;
SQ
Query Match 85.6%; Score 15.4; DB 12; Length 1664;
Best Local Similarity 94.1%; Pred. No. 3.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGATCTCTCCA 17
Db 560 CGACTGTGATCTCTCCA 576
RESULT 33
AAV55741
ID AAV55741 standard; cDNA; 1755 BP.
XX
XX AAV55741;
XX
XX 23-MAR-1999 (first entry)
DT
XX Human secreted protein clone as294_3 coding sequence.
DE
XX
XX Secreted protein; human; nutritional supplements; immune stimulant;
XX KW immune suppressor; haematopoiesis regulator; tissue growth; chemotaxis;
XX KW activin/inhibin; chemokinesis; haemostasis; thrombolysis;
XX KW receptor/ligand activity; anti-inflammatory; tumour inhibitor;
XX KW cadherin/tumour invasion suppressor; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX FT CDS 266..1654
XX FT /*tag= a
XX
XX MO9855614-A2.
XX
XX 10-DEC-1998.
XX
XX 01-JUN-1998; 98WO-US011210.
XX
XX 04-JUN-1997; 97US-00868696.
XX PR 04-JUN-1997; 97US-00868697.
XX PR 04-JUN-1997; 97US-00868698.
XX PR 04-JUN-1997; 97US-00868699.
XX PR 04-JUN-1997; 97US-00868900.
XX PR 04-JUN-1997; 97US-00869191.
XX PR 04-JUN-1997; 97US-00869192.
XX PR 04-JUN-1997; 97US-00869193.
XX PR 04-JUN-1997; 97US-00869194.
XX PR 29-MAY-1998; 98US-0087252P.
XX
XX (GEMV ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Treacy M, Spaulding V;
XX PI Agostino MJ, Howes SR, Fehnel K;
XX
XX WPI; 1999-059912/05.
XX P-PSDB; AAW73623.
XX

```

PT New polynucleotides encoding secreted human proteins - derived from human
 PT foetal brain, adult brain, foetal kidney, adult ovary, adult retina,
 PT adult placenta or adult uterus cDNA libraries.

XX Claim 1; Page 71-72; 127pp; English.

CC This sequence encodes a human secreted protein of the invention. This DNA
 CC sequence was isolated from a human foetal brain cDNA library, and was
 CC designated clone as294.3. The DNAs and proteins are predicted to have
 CC biological activities which would make them suitable for treating,
 CC preventing or ameliorating medical conditions in humans and animals,
 CC although no supporting data is given. Suggested activities include
 CC nutritional sources or supplements, immune stimulating or suppressing
 CC activity, haematopoiesis regulating activity, tissue growth activity,
 CC activ/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, and tumour
 CC inhibition activity. The DNAs are also stated to be useful for gene
 CC therapy. A host cell transfected with the DNA, or its subfragments and
 CC variants is useful for recombinant production of the human secreted
 CC protein clones

XX Sequence 1755 BP; 384 A; 519 C; 452 G; 399 T; 0 U; 1 Other;

Query Match 85.6%; Score 15.4; DB 2; Length 1755;
 Best Local Similarity 94.1%; Pred. No. 3.4e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGATTCCTCCA 17
 DB 673 CGACTGTGATTCCTCCA 689

RESULT 34

ID ABO92067 standard; cDNA; 1755 BP.

XX ABO92067;

DT 04-OCT-2002 (first entry)

XX Human polynucleotide SEQ ID NO 64.

XX Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic;
 KW antinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
 KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
 KW antidiabetic; antidiabetic; antidiabetic; antidiabetic;
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KW stem cell; growth factor; nervous system disease; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;
 KW multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.

XX Homo sapiens.

OS US2002065394-A1.

XX 30-MAY-2002.

XX 22-DEC-2000; 2000US-00745763.

XX 18-MAR-1998; 98US-00040963.

XX (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVALL) LAVALLIE E R.
 PA (COLL) COLLINS-RACIE L A.
 PA (EVAN) EVANS C.
 PA (MERBER) MERBERG D.
 PA (TREAC) TREACY M.
 PA (SPAUL) SPAULDING V.

PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 XX WPI; 2002-582343/62.
 DR P-PSDB; ABP61854.

XX Novel secreted or transmembrane protein and polynucleotide encoding the
 CC protein, useful for diagnosis and treatment of neurological disorders,
 CC cancer, autoimmune diseases, bone disorders and lung or liver fibrosis.
 CC Claim 227; Page 219-220; 284pp; English.

XX The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (cDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (II) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, anti-inflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (II) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polynucleotide of the invention

XX Sequence 1755 BP; 384 A; 519 C; 452 G; 399 T; 0 U; 1 Other;

Query Match 85.6%; Score 15.4; DB 6; Length 1755;
 Best Local Similarity 94.1%; Pred. No. 3.4e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGACTGTGATTCCTCCA 17
 DB 673 CGACTGTGATTCCTCCA 689

RESULT 35

ID AAF21834 standard; DNA; 1892 BP.

XX AAF21834;

DT 27-MAR-2001 (first entry)

XX Human breast and ovarian cancer associated antigen gene SEQ ID 221.

XX Human; breast cancer; ovarian cancer; cytosolic; immunosuppressive;
 KW neurotropic; neuroprotective; antiviral; antiarthritic; hepatotropic;
 KW antidiabetic; antinflammatory; antidiabetic; antidiabetic;
 KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease; ds.

XX Homo sapiens.
 OS
 XX
 PN WO200055173-A1.
 XX
 PD 21-SEP-2000.
 XX
 PP 08-MAR-2000; 2000WO-US005881.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX WPI; 2000-611515/58.
 DR P-PSDB; AAB58931.
 XX
 PT New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention, treatment
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
 PT neurological diseases.
 XX
 PS Claim 1; Page 647-648; 1299pp; English.
 XX
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic, immunosuppressive, nocrotropic,
 CC neuroprotective, antiviral, antiallergic, hepatotropic, antidiabetic;
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
 CC protein sequences are used in the diagnosis of cancer, particularly
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
 CC and agonists may also be used in the diagnosis, prevention and treatment
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC cardiovascular disorders such as myocardial ischaemia; wound healing;
 CC neurological diseases such as cerebral anoxia and epilepsy; and
 CC infectious diseases
 XX
 SQ Sequence 1892 BP; 441 A; 553 C; 478 G; 417 T; 0 U; 3 Other;
 Query Match 85.6%; Score 15.4; DB 3; Length 1892;
 Best Local Similarity 94.1%; Pred. No. 3.4e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCTCA 17
 |||||
 DB 763 CGACTGTGATCTCTCA 779

RESULT 36
 ABL19347/c
 ID ABL19347 standard; DNA; 2148 BP.
 XX
 AC ABL19347;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9514.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.

PD 27-SEP-2001.
 XX
 PP 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PERE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 9514; 21pp + Sequence listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutic and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins (AAB57737-
 CC AAB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2148 BP; 534 A; 559 C; 506 G; 549 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 4; Length 2148;
 Best Local Similarity 94.1%; Pred. No. 3.4e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGATCTCTCA 17
 |||||
 DB 444 CGACTGTGATCTCTCA 428

RESULT 37
 ABL15659/c
 ID ABL15659 standard; CDNA; 2243 BP.
 XX
 AC ABL15659;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 41459.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 EN
 XX
 PD 27-SEP-2001.
 XX
 PP 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PERE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR P-PSDB; AAB71556.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more

PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232396P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0234400P.
PR 14-SEP-2000; 2000US-0234401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239356P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251988P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI, 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides.
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclousure; SEQ ID NO 29980; 3071pp + Sequence listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosratic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK4703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX SQ Sequence 4357 BP; 1312 A; 834 C; 950 G; 1261 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 4; Length 4357;
Best Local Similarity 94.1%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GACTGTGAATCCTCCAT 18
DB 3193 GACTGTGAATCCTCCAT 3209
RESULT 40
ABLI9346/C
ID ABLI9346 standard; DNA; 4532 BP.
XX
XX ABLI9346;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9511.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ds.
 XX
 XX Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 XX 27-SEP-2001.
 PD
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 PA (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 CC
 PS Claim 1; SEQ ID NO 9511; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 4532 BP; 1263 A; 1022 C; 975 G; 1272 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 4; Length 4532;
 Best Local Similarity 94.1%; Pred. No. 3.7e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGACTGGAATCCCA 17
 Db 1649 CGACAGTGAATCTCCA 1633
 RESULT 41
 ABL15658/c
 ID ABL15658 standard; cDNA; 6446 BP.
 XX
 AC ABL15658;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 41456.
 XX
 KM Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 XX 27-SEP-2001.
 PD
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 PR

XX
 PA (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR P-PSDB; AB571555.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 CC
 PS Claim 1; SEQ ID NO 41456; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 6446 BP; 1819 A; 1415 C; 1394 G; 1818 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 4; Length 6446;
 Best Local Similarity 94.1%; Pred. No. 3.8e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGACTGGAATCCCA 17
 Db 3563 CGACAGTGAATCTCCA 3547
 RESULT 42
 ADO47193
 ID ADO47193 standard; DNA; 200000 BP.
 XX
 AC ADO47193;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE DNA sequence of a human immunoglobulin heavy chain variable region.
 XX
 KM immunoglobulin; B cell; germline; region V; region D; region J;
 KM recombination-promoting protein; recombination activating gene 1; RAG-1;
 KM RAG-2; ss; terminal deoxynucleotidyltransferase; Tdt; V(D)J recombinase.
 XX
 OS Homo sapiens.
 XX
 PN WO2004029249-A1.
 XX
 PD 08-APR-2004.
 XX
 PF 30-SEP-2003; 2003WO-AU001286.
 XX
 PR 30-SEP-2002; 2002US-0415024P.
 XX
 PA (CENT-) CENTENARY INST CANCER MEDICINE & CELL BI.
 XX
 PI Jolly C;
 XX
 DR WPI; 2004-316126/29.
 XX
 XX New vector comprising one or more immunoglobulin regions selected from V,
 PT D, and J regions, a 5' flanking region, and a 3' flanking region, useful
 PT for reverting cell lines to a pro-B cell-like state or to a germline-like
 PT state.
 XX
 PS Example 1; Page 234-283; 382bp; English.
 XX

CC The specification describes a method for the preparation of
 CC immunoglobulins. The method uses a vector for reverting cell lines to a
 CC pro-B cell-like state or to a germline-like state, by replacing one or
 CC more of the immunoglobulin regions V_H, D_H, and J of the rearranged
 CC immunoglobulin gene with the germ-line or unrearranged V_H, D or J regions.
 CC The vector can comprise a polynucleotide sequence encoding a
 CC recombination-promoting protein, selected from recombination activating
 CC gene 1 (RAG-1), RAG-2, terminal deoxynucleotidyltransferase (Tdt). These
 CC proteins collectively constitute a V(D)J recombinase. The method is
 CC useful for the preparation of immunoglobulin heavy and/or light chains
 CC and their fragments. The present sequence encodes a human immunoglobulin
 CC heavy chain variable region, and is used as a template to construct
 CC vectors for use in the method of the invention.

XX Sequence 20000 BP, 59140 A, 40391 C, 40764 G, 59705 T, 0 U, 0 Other;

Query Match 85.6%; Score 15.4; DB 12; Length 200000;
 Best Local Similarity 94.1%; Pred. No. 5e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACTGTGATCTCCAT 18
 DB 107278 GACTGAGATCTCCAT 107294

RESULT 43

ACAG1261 ID ACAG1261 standard; DNA, 18 BP.

XX ACAG1261;

XX 11-AUG-2003 (first entry)

XX Rat p53 antisense oligonucleotide.

XX Rat; ss; antisense; p53; bioactive agent delivery.

XX Rattus norvegicus.

XX US2003026831-A1.

XX 06-FEB-2003.

XX 22-APR-2002; 2002US-00131786.

XX 20-APR-2001; 2001US-0285337P.

XX (LAKK/) LAKKARAJU A.

XX (DUBI/) DUBINSKY J M.

XX (LOWN/) LOW W.

XX (RAHM/) RAHMAN Y.

XX Lakkaraju A, Dubinsky JM, Low W, Rahman Y;

XX WPI; 2003-439031/41.

XX Claim 21, Page 5; 56pp; English.

XX The invention relates to a pharmaceutical composition comprising an
 CC anionic liposome, a bioactive agent, a cation and/or a buffer. The
 CC composition is useful for delivering a bioactive agent to a target e.g.
 CC cell of an organ (e.g. brain, central nervous system, peripheral nervous
 CC system, liver, lung, larynx, bone marrow, spleen, kidney, lymphatic
 CC system, hematopoietic system, gastric mucosa, small intestine, large
 CC intestine, gall bladder, pancreas, salivary gland, testes, ovary, cervix,
 CC uterus, muscle, skin, thyroid gland, parathyroid gland, adrenal gland,
 CC connective tissue, chondroid tissue, blood vessel, macrophage, pleura,
 CC placenta, a tumour or a growth), non-dividing cells, neuronal cells,
 CC hippocampal neuronal cells, cell that expresses a receptor belonging to

CC the low-density lipoprotein (LDL) gene family, cell that possesses (LDL)
 CC receptor-related protein (LRP) receptor, cell that possesses an endocytic
 CC (LDL) lipoprotein receptor-related protein receptor, cell that possesses
 CC a receptor that is expressed in mammalian central nervous system (CNS),
 CC pluripotent cell, stem cell, embryonic stem cells and bone marrow
 CC derived stem cells. The present sequence represents the rat p53 antisense
 CC oligonucleotide

XX Sequence 18 BP, 4 A, 6 C, 3 G, 5 T, 0 U, 0 Other;

Query Match 83.3%; Score 15; DB 10; Length 18;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CTGTGATCTCCAT 18
 DB 1 CTGTGATCTCCAT 15

RESULT 44

AA297144/C ID AA297144 standard; cDNA, 1024 BP.

XX AA297144;

XX 18-APR-2000 (first entry)

XX Human prostate cancer differentially expressed gene #5.

XX Prostate cancer specific gene; cancer; tumour progression; diagnose;

XX hyperproliferative cell growth; prostatic disorder; treatment;

XX metastatic prostate cancer; benign prostate hyperplasia; BPH; ss.

XX Homo sapiens.

XX WO964594-A2.

XX 16-DEC-1999.

XX 10-JUN-1999; 99WO-US013181.

XX 11-JUN-1998; 98US-0088877P.

XX 09-JUN-1999; 99US-00284475.

XX (CHIR) CHIRON CORP.

XX Astel JH, Carroll B, Endege WO, Ford DM, Monahan JB, Schlegel R;

XX Steimann KB, Zhang J;

XX WPI; 2000-116541/10.

XX Claim 12, Page 72; 212pp; English.

XX New isolated prostate cancer specific nucleic acids, used to develop

XX products for the diagnosis and treatment of cancer.

XX This sequence represents a prostate cancer specific nucleic acid

XX sequence. The invention relates to a method for diagnosing cancer, tumour

XX progression, hyperproliferative cell growth or accompanying biological

XX and physical manifestations. The method involves contacting the

XX biological sample with a probe that comprises a sequence capable of

XX hybridising to any of the 339 nucleotide sequences given in the

XX specification (see AA297140-297478) and detecting duplex formation. The

XX products and methods of the invention can be used for the diagnosis,

XX prognosis, and treatment of cancer, tumour progression,

XX hyperproliferative cell growth, and accompanying physical and biological

XX manifestations. They can be used particularly for prostatic disorders

XX such as metastatic prostate cancer, localised prostate cancer, or benign

XX prostate hyperplasia (BPH)

XX Sequence 1024 BP, 230 A, 217 C, 251 G, 191 T, 0 U, 135 Other;

XX Query Match 83.3%; Score 15; DB 3; Length 1024;

Best Local Similarity 93.8%; Pred. No. 5.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ACTGTGAATCCTCCAT 18
Db 574 ACTGTGAATCCTCCAT 559

RESULT 45

ABN66254/C
ID ABN66254 standard; DNA; 1065 BP.

ABN66254;

01-JUL-2002 (first entry)

Streptococcus polynucleotide SEQ ID NO 421.

Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.

Streptococcus agalactiae.

MO200234771-A2.

02-MAY-2002.

29-OCT-2001; 2001WO-GB004789.

27-OCT-2000; 2000GB-00026333.

24-NOV-2000; 2000GB-00028727.

07-MAR-2001; 2001GB-00005640.

(CHIR-) CHIRON SPA.

(GENO-) INST GENOMIC RES.

Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;

Tetzelin H;

WPI: 2002-352536/38.

P-PSDB; ABP25623.

Claim 7; Page 3197; 4525BP; English.

The invention relates to a protein (ABP25413-ABP30895) from group B

streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

the specification. The proteins have antibacterial and antiinflammatory

activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and

antibodies that bind (1) are used in the manufacture of medicaments for

the treatment or prevention of infection or disease caused by

Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

Nucleic acids encoding (1) are used to detect Streptococcus in a

biological sample. (1) is used to determine whether a compound binds to

(1). A composition comprising (1) or a nucleic acid encoding (1), may be

used as a vaccine or diagnostic composition. The disease caused by

Streptococcus that is prevented or treated may be meningitis. Nucleic

acid encoding (1) may be used to recombinantly produce (1) and may be

used in gene therapy. Antibodies to (1) are used for affinity

chromatography, immunoassays, and distinguishing/identifying

Qy 4 CTGTGAATCCTCCAT 18
Db 436 CTGTGAATCCTCCAT 422

RESULT 46

ABN71527_01
Continuation (2 of 22) of ABN71527 from base 100001 (Streptococcus polynucleotide SEQ ID

WP Sequence split into 22 fragments LOCUS ABN71527 Accession ABN71527

Fragment Name

WP ABN71527_00

WP ABN71527_01

WP ABN71527_02

WP ABN71527_03

WP ABN71527_04

WP ABN71527_05

WP ABN71527_06

WP ABN71527_07

WP ABN71527_08

WP ABN71527_09

WP ABN71527_10

WP ABN71527_11

WP ABN71527_12

WP ABN71527_13

WP ABN71527_14

WP ABN71527_15

WP ABN71527_16

WP ABN71527_17

WP ABN71527_18

WP ABN71527_19

WP ABN71527_20

WP ABN71527_21

WP

WP

WP

WP

WP

WP

WP

WP

WP

WP

Query Match 83.3%; Score 15; DB 6; Length 110000;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CTGTGAATCCTCCAT 18
Db 66795 CTGTGAATCCTCCAT 66809

RESULT 47

ADL71108
ID ADL71108 standard; DNA; 93 BP.

ADL71108;

01-JUL-2004 (first entry)

p53 hairpin oligo p53-2.

ss; p53; hairpin; stem cell; short hairpin RNA; shRNA; anti-HIV;

cytotoxic; cell therapy; HIV; cancer.

Unidentified.

WO2004029219-A2.

08-APR-2004.

29-SEP-2003; 2003WO-US030901.

27-SEP-2002; 2002US-0414605P.

(COLD-) COLD SPRING HARBOR LAB.

Fridman J, Hannon GJ, Hemann M, Lowe SW, Paddison PJ, Zilfou J;

Dickins R, Carmell MA, Rosenquist TA;

WPI: 2004-316101/29.

Introducing into a subject stem cells with partial or complete loss of

PT function of a target gene comprises introducing nucleic acids encoding
PT short hairpin RNAs into stem cells and introducing the transfected stem
PT cells into the subject.
XX
PS Example 1, Page 37; 93pp; English.
XX
CC The invention relates to a novel method for introducing into a subject a
CC population of stem cells having partial or complete loss of function of a
CC target gene by introducing a nucleic acid construct encoding a short
CC hairpin RNA (shRNA) into stem cells to generate transfected stem cells
CC and introducing the transfected stem cells into the subject. The
CC transfected stem cells propagate within the subject and retain partial to
CC complete loss of function of the target gene. The invention has anti-HIV
CC and cytostatic activity, and may have a use in cell therapy. The
CC composition may be administered intravenously or intraarterially, or into
CC the liver (via portal vein injection), the peritoneal cavity, the kidney
CC capsule or the bone marrow. The composition and methods are useful for
CC performing RNA interference in stem cells and for treating diseases (e.g.
CC HIV or cancer) associated with the expression of a target gene in a
CC population of cells. The present sequence represents a shRNA nucleic acid
CC of the invention.
SQ Sequence 93 BP; 26 A; 27 C; 19 G; 21 T; 0 U; 0 Other;
XX
Query Match 82.2%; Score 14.8; DB 12; Length 93;
Best Local Similarity 88.9%; Pred. No. 5.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGACTGTGATCTCTCCAT 18
13 CGACTGCGACTCTCTCCAT 30
Db 13 CGACTGCGACTCTCTCCAT 30
RESULT 48
ADQ05153
ID ADQ05153 standard; DNA; 281 BP.
XX
AC ADQ05153;
XX
DT 09-SEP-2004 (first entry)
XX
DE Soybean homeobox transcription factor seqid 1999.
XX
KW maize; soybean; Arabidopsis thaliana; transcription factor; homeobox;
KW HbH; leucine zipper; zinc finger; transformed plant; plant;
KW metabolic pathway; mutation detection; polymorphism; plant trait;
KW genome mapping; gene identification; gene analysis; plant breeding;
KW transgenic; ds.
XX
OS Glycine max.
XX
PN US2004123339-A1.
XX
PD 24-JUN-2004.
XX
PF 06-AUG-2001; 2001US-00922293.
XX
PR 24-NOV-1997; 97US-0067000P.
PR 09-DEC-1997; 97US-0069472P.
PR 13-JAN-1998; 98US-0071479P.
PR 10-FEB-1998; 98US-0074201P.
PR 10-FEB-1998; 98US-0074280P.
PR 10-FEB-1998; 98US-0074281P.
PR 10-FEB-1998; 98US-0074282P.
PR 12-FEB-1998; 98US-0074565P.
PR 12-FEB-1998; 98US-0074566P.
PR 12-FEB-1998; 98US-0074567P.
PR 12-FEB-1998; 98US-0074789P.
PR 19-FEB-1998; 98US-0075459P.
PR 19-FEB-1998; 98US-0075460P.
PR 19-FEB-1998; 98US-0075461P.
PR 19-FEB-1998; 98US-0075462P.
PR 19-FEB-1998; 98US-0075463P.

PR 19-FEB-1998; 98US-0075464P.
PR 09-MAR-1998; 98US-0077229P.
PR 09-MAR-1998; 98US-0077230P.
PR 09-MAR-1998; 98US-0077231P.
PR 18-MAR-1998; 98US-0078368P.
PR 07-APR-1998; 98US-0080844P.
PR 27-APR-1998; 98US-0083067P.
PR 29-APR-1998; 98US-0083366P.
PR 29-APR-1998; 98US-0083367P.
PR 29-APR-1998; 98US-0083368P.
PR 29-APR-1998; 98US-0083369P.
PR 13-MAY-1998; 98US-0085222P.
PR 13-MAY-1998; 98US-0085223P.
PR 13-MAY-1998; 98US-0085224P.
PR 15-MAY-1998; 98US-0085533P.
PR 21-MAY-1998; 98US-0086183P.
PR 21-MAY-1998; 98US-0086184P.
PR 21-MAY-1998; 98US-0086185P.
PR 21-MAY-1998; 98US-0086186P.
PR 21-MAY-1998; 98US-0086187P.
PR 21-MAY-1998; 98US-0086188P.
PR 16-JUN-1998; 98US-0089524P.
PR 18-JUN-1998; 98US-0089793P.
PR 18-JUN-1998; 98US-0089806P.
PR 18-JUN-1998; 98US-0089807P.
PR 18-JUN-1998; 98US-0089808P.
PR 18-JUN-1998; 98US-0089810P.
PR 18-JUN-1998; 98US-0089811P.
PR 18-JUN-1998; 98US-0089812P.
PR 18-JUN-1998; 98US-0089813P.
PR 18-JUN-1998; 98US-0089814P.
PR 30-JUN-1998; 98US-0091247P.
PR 30-JUN-1998; 98US-0091405P.
PR 09-SEP-1998; 98US-0093667P.
PR 09-SEP-1998; 98US-0093668P.
PR 09-SEP-1998; 98US-0093670P.
PR 09-SEP-1998; 98US-0093672P.
PR 16-SEP-1998; 98US-0100672P.
PR 16-SEP-1998; 98US-0100673P.
PR 16-SEP-1998; 98US-0100674P.
PR 17-SEP-1998; 98US-0100963P.
PR 21-SEP-1998; 98US-0101130P.
PR 21-SEP-1998; 98US-0101131P.
PR 21-SEP-1998; 98US-0101132P.
PR 22-SEP-1998; 98US-0101343P.
PR 22-SEP-1998; 98US-0101344P.
PR 22-SEP-1998; 98US-0101347P.
PR 22-SEP-1998; 98US-0101508P.
PR 25-SEP-1998; 98US-0101707P.
PR 13-OCT-1998; 98US-0104124P.
PR 13-OCT-1998; 98US-0104126P.
PR 13-OCT-1998; 98US-0104127P.
PR 13-OCT-1998; 98US-0104128P.
PR 18-NOV-1998; 98US-0108996P.
PR 19-NOV-1998; 98US-0109018P.
PR 24-NOV-1998; 98US-00199129.
PR 08-DEC-1998; 98US-002109297.
PR 11-DEC-1998; 98US-0111981P.
PR 22-DEC-1998; 98US-0113224P.
PR 12-JAN-1999; 99US-00229413.
XX
PA (CONN/) CONNER T W.
PA (HECK/) HECK G R.
PA (LIU/) LIU J.
XX
PI Conner TW, Heck GR, Liu J;
XX
DR WPI; 2004-468202/44.
XX
PT New substantially purified nucleic acid molecule that encodes a maize,
PT soybean or Arabidopsis thaliana transcription factor or its fragment,
PT useful for genome mapping, gene identification and analysis or plant
PT breeding.

XX Claim 2; SEQ ID NO 1999; 140pp; English.

PS The invention describes a substantially purified nucleic acid molecule

CC that encodes a maize, soybean or Arabidopsis thaliana transcription

CC factor or its fragment, where the maize or soybean transcription factor

CC is homeobox, HbH, leucine zipper, zinc finger, or other transcription

CC factor. Also described are: a substantially purified maize or soybean

CC transcription factor or its fragment defined above; a substantially

CC purified antibody or its fragment which is capable of specifically

CC binding to the transcription factor or its fragment above; a transformed

CC plant; a method for determining a level or pattern in a plant cell of a

CC transcription factor in a plant metabolic pathway; a method of

CC determining a mutation in a plant; a method of producing a plant

CC containing an overexpressed protein or reduced levels of plant

CC transcription factor; a method of determining an association between a

CC polymorphism and a plant trait; and a method of isolating a nucleic acid

CC that encodes a plant transcription factor or its fragment. The nucleic

CC acid molecules, proteins and their fragments are useful for genome

CC mapping, gene identification and analysis, plant breeding, preparation of

CC constructs for use in plant gene expression and transgenic plants. The

CC nucleic acid molecules are useful as markers or probes. This sequence

CC represents a soybean homeobox transcription factor polynucleotide.

XX Sequence 281 BP; 87 A; 63 C; 63 G; 68 T; 0 U; 0 Other;

SQ

Query Match 82.2%; Score 14.8; DB 12; Length 281;

Best Local Similarity 88.9%; Pred. No. 6.1e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18

Db 258 CGAGTGTGATCCTCCAT 275

RESULT 49

ABL66464

ID ABL66464 standard; DNA; 417 BP.

XX

AC ABL66464;

XX

DT 15-MAY-2002 (first entry)

XX

XX Lung cancer related gene sequence SEQ ID NO:4801.

XX

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KW cytoskeletal; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;

KW gene; ds.

XX

OS Homo sapiens.

XX

PN WO200194629-A2.

XX

PD 13-DEC-2001.

XX

PF 30-MAY-2001; 2001WO-US010838.

XX

XX 05-JUN-2000; 2000US-0209473P.

PR 05-JUN-2000; 2000US-0209531P.

XX

XX 18-SEP-2000; 2000US-0233133P.

PR 18-SEP-2000; 2000US-0233617P.

XX

XX 20-SEP-2000; 2000US-0234009P.

PR 20-SEP-2000; 2000US-0234034P.

XX

XX 20-SEP-2000; 2000US-0234052P.

PR 22-SEP-2000; 2000US-0234509P.

XX

XX 22-SEP-2000; 2000US-0234567P.

PR 22-SEP-2000; 2000US-0234923P.

XX

XX 25-SEP-2000; 2000US-0234924P.

PR 25-SEP-2000; 2000US-0235077P.

XX

XX 25-SEP-2000; 2000US-0235082P.

PR 25-SEP-2000; 2000US-0235134P.

XX

XX 25-SEP-2000; 2000US-0235280P.

PR 26-SEP-2000; 2000US-0235637P.

PR 26-SEP-2000; 2000US-0235638P.

XX

PR 27-SEP-2000; 2000US-0235711P.

PR 27-SEP-2000; 2000US-0235720P.

XX

PR 27-SEP-2000; 2000US-0235840P.

PR 27-SEP-2000; 2000US-0235863P.

XX

PR 28-SEP-2000; 2000US-0236028P.

PR 28-SEP-2000; 2000US-0236032P.

XX

PR 28-SEP-2000; 2000US-0236033P.

PR 28-SEP-2000; 2000US-0236034P.

XX

PR 28-SEP-2000; 2000US-0236111P.

PR 28-SEP-2000; 2000US-0236842P.

XX

PR 29-SEP-2000; 2000US-0236881P.

PR 29-SEP-2000; 2000US-0237172P.

XX

PR 02-OCT-2000; 2000US-0237173P.

PR 02-OCT-2000; 2000US-0237278P.

XX

PR 02-OCT-2000; 2000US-0237294P.

PR 02-OCT-2000; 2000US-0237295P.

XX

PR 02-OCT-2000; 2000US-0237316P.

PR 03-OCT-2000; 2000US-0237425P.

XX

PR 03-OCT-2000; 2000US-0237598P.

PR 03-OCT-2000; 2000US-0237604P.

XX

PR 03-OCT-2000; 2000US-0237606P.

PR 03-OCT-2000; 2000US-0237608P.

XX

PR 01-NOV-2000; 2000US-0244867P.

PR 01-NOV-2000; 2000US-0245084P.

XX

PA (AVAL-) AVALON PHARM.

XX

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

PI Soppet DR, Weaver Z;

XX

XX WPI; 2002-188264/24.

DR

XX

XX Screening for anti-neoplastic agent involves exposing cells to a chemical

PT agent to be tested for anti-neoplastic activity, and determining a change

PT in expression of a gene of a signature gene set.

XX

PS Claim 1; SEQ ID NO 4801; 44pp; English.

XX

XX The present invention describes a method (M1) for screening for an anti-

CC neoplastic agent. The method involves exposing cells to a chemical agent

CC to be tested for anti-neoplastic activity, determining a change in

CC expression of at least one gene (I) of a signature gene set, where (I)

CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664

CC to ABL70110), or is at least 95% identical to (S), where a change in

CC expression is indicative of anti-neoplastic activity. (I) has cytoskeletal

CC activity and can be used in gene therapy. M1 can be used for screening an

CC anti-neoplastic agent, and can be used for producing a product which is

CC the data collected with respect to the anti-neoplastic agent as a result

CC of M1, and the data is sufficient to convey the chemical structure and/or

CC properties of the agent. M1 can be used in the treatment of cancer such

CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,

CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell

CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous

CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'

CC tumour

XX

SQ Sequence 417 BP; 92 A; 130 C; 93 G; 101 T; 0 U; 1 Other;

XX

Query Match 82.2%; Score 14.8; DB 6; Length 417;

Best Local Similarity 88.9%; Pred. No. 6.3e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18

Db 278 CGACTGTGATCCTCCAT 295

RESULT 50

ABL65765

ID ABL65765 standard; DNA; 417 BP.

XX ABL65765;
AC
XX 15-MAY-2002 (first entry)
DT
XX Lung cancer related gene sequence SEQ ID NO:4102.
DE
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; anti-neoplastic; Wilms tumour; adenocarcinoma;
KW gene; ds.
OS Homo sapiens.
XX
XX WO200194629-A2.
PN
XX 13-DEC-2001.
PD
XX 30-MAY-2001; 2001WO-US010838.
PF
XX 05-JUN-2000; 2000US-0209473P.
PR
XX 05-JUN-2000; 2000US-0209531P.
PR
XX 18-SEP-2000; 2000US-0231133P.
PR
XX 18-SEP-2000; 2000US-0233617P.
PR
XX 20-SEP-2000; 2000US-0234009P.
PR
XX 20-SEP-2000; 2000US-0234034P.
PR
XX 20-SEP-2000; 2000US-0234052P.
PR
XX 22-SEP-2000; 2000US-0234509P.
PR
XX 22-SEP-2000; 2000US-0234567P.
PR
XX 25-SEP-2000; 2000US-0234923P.
PR
XX 25-SEP-2000; 2000US-0234924P.
PR
XX 25-SEP-2000; 2000US-0235077P.
PR
XX 25-SEP-2000; 2000US-0235082P.
PR
XX 25-SEP-2000; 2000US-0235134P.
PR
XX 25-SEP-2000; 2000US-0235280P.
PR
XX 26-SEP-2000; 2000US-0235637P.
PR
XX 26-SEP-2000; 2000US-0235638P.
PR
XX 27-SEP-2000; 2000US-0235711P.
PR
XX 27-SEP-2000; 2000US-0235720P.
PR
XX 27-SEP-2000; 2000US-0235840P.
PR
XX 27-SEP-2000; 2000US-0235863P.
PR
XX 28-SEP-2000; 2000US-0236028P.
PR
XX 28-SEP-2000; 2000US-0236033P.
PR
XX 28-SEP-2000; 2000US-0236033P.
PR
XX 28-SEP-2000; 2000US-0236034P.
PR
XX 28-SEP-2000; 2000US-0236109P.
PR
XX 28-SEP-2000; 2000US-0236111P.
PR
XX 29-SEP-2000; 2000US-0236842P.
PR
XX 29-SEP-2000; 2000US-0236891P.
PR
XX 02-OCT-2000; 2000US-0237172P.
PR
XX 02-OCT-2000; 2000US-0237173P.
PR
XX 02-OCT-2000; 2000US-0237278P.
PR
XX 02-OCT-2000; 2000US-0237294P.
PR
XX 02-OCT-2000; 2000US-0237295P.
PR
XX 02-OCT-2000; 2000US-0237316P.
PR
XX 03-OCT-2000; 2000US-0237425P.
PR
XX 03-OCT-2000; 2000US-0237598P.
PR
XX 03-OCT-2000; 2000US-0237604P.
PR
XX 03-OCT-2000; 2000US-0237606P.
PR
XX 03-OCT-2000; 2000US-0237608P.
PR
XX 01-NOV-2000; 2000US-0244867P.
PR
XX 01-NOV-2000; 2000US-0245084P.
PR
XX (AVAL-) AVALON PHARM.
PA
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
PI
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.

XX Claim 1; SEQ ID NO 4102; 44pp; English.
PS
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
CC tumour
XX
XX Sequence 417 BP; 92 A; 130 C; 93 G; 101 T; 0 U; 1 Other;
SQ
XX
XX Query Match 82.2%; Score 14.8; DB 6; Length 417;
Best Local Similarity 88.9%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 CGACTGTGATTCCTCCAT 18
Db 278 CGACTGTGATTCACCAT 295
|||||
|||||

Search completed: January 14, 2005, 16:13:50
Job time : 359.263 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: January 14, 2005, 10:01:39 ; Search time 913.263 Seconds
(without alignments)
932.058 Million cell updates/sec

Title: US-09-578-453-1

Perfect score: 18

Sequence: 1 CGACTGCAATCCTCCAT 18

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 500 summaries

Database :

GenEmbl:*
1: gb_ha:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_rts:*
12: gb_by:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	199	10	RATP53TS02
2	18	100.0	513	10	AY009504
3	18	100.0	1019	10	AF209191
4	18	100.0	1176	10	RNU90328
5	18	100.0	1627	6	AX401815
6	18	100.0	1627	6	AX827834
7	18	100.0	1627	10	RNP53
8	18	100.0	9414	12	AY297714
9	18	100.0	165316	2	AC119115
10	18	100.0	192326	2	AC134317
11	18	100.0	225077	2	AC136563
12	17	94.4	129489	10	AL713860
13	17	94.4	150816	2	AC119532
14	17	94.4	181075	10	CNS071P2
15	17	94.4	229363	2	AC094940
16	17	94.4	255706	2	AC094186
17	17	94.4	265074	2	AC103024
18	17	94.4	267779	2	AC112826
19	17	94.4	275838	2	AL627428

20	17	94.4	328868	2	AC118955
21	16.4	91.1	84	10	MMANR02
22	16.4	91.1	390	6	AX435258
23	16.4	91.1	836	3	AB038502
24	16.4	91.1	1096	10	MMU297973
25	16.4	91.1	1173	6	AR052877
26	16.4	91.1	1173	6	AR127895
27	16.4	91.1	1173	6	AR288189
28	16.4	91.1	1176	10	AY044188
29	16.4	91.1	1183	10	AF161020
30	16.4	91.1	1208	10	AF051368
31	16.4	91.1	1241	10	MUSP53B
32	16.4	91.1	1285	10	MUSP53A
33	16.4	91.1	1322	10	MUSP53C
34	16.4	91.1	1360	10	AB017815
35	16.4	91.1	1377	10	MMP53
36	16.4	91.1	1409	10	AF151353
37	16.4	91.1	1429	10	AB021961
38	16.4	91.1	1457	10	AB020317
39	16.4	91.1	1746	10	AY212017
40	16.4	91.1	1773	10	MMP53R
41	16.4	91.1	1782	10	BC005448
42	16.4	91.1	2130	10	K02110
43	16.4	91.1	2130	10	K02110
44	16.4	91.1	86146	2	AC023805
45	16.4	91.1	116898	10	AL731687
46	16.4	91.1	150937	9	AL445243
47	16.4	91.1	163447	2	AC102741
48	16.4	91.1	172568	2	BX324147
49	16.4	91.1	201602	2	AC074146
50	16.4	91.1	214262	2	AC148732
51	16.4	91.1	229486	2	AC074149
52	16.4	91.1	263036	2	AC105597
53	16.4	91.1	283735	2	AC098078
54	16.4	88.9	3890	14	RNMR81
55	16	88.9	104422	9	AC009656
56	16	88.9	139166	9	AC011750
57	16	88.9	140680	8	AF131112
58	16	88.9	146836	9	AC019317
59	16	88.9	151660	2	AC027450
60	16	88.9	154827	9	AL592221
61	16	88.9	170950	14	AY510475
62	16	88.9	174077	14	AF147806
63	16	88.9	177874	14	AF243438
64	16	88.9	180105	9	AC015756
65	16	88.9	182549	9	AC061999
66	16	88.9	184254	9	AC131935
67	16	88.9	197379	2	AC119278
68	16	88.9	206511	2	AC137798
69	16	88.9	211009	8	AP326781
70	16	88.9	212270	10	AC108811
71	16	88.9	238305	2	AC139741
72	15.4	85.6	296	9	AF027558
73	15.4	85.6	315	10	AF196265
74	15.4	85.6	319	10	AF196264
75	15.4	85.6	320	10	AF196260
76	15.4	85.6	325	10	AF196262
77	15.4	85.6	327	10	AF196263
78	15.4	85.6	347	10	AF196261
79	15.4	85.6	351	10	AF196255
80	15.4	85.6	352	10	AF196259
81	15.4	85.6	367	10	RNG101V3
82	15.4	85.6	367	10	RNMBEVA2
83	15.4	85.6	396	6	A42904
84	15.4	85.6	407	6	AR271438
85	15.4	85.6	420	10	MMU07878
86	15.4	85.6	430	10	RNU76836
87	15.4	85.6	525	6	AX867656
88	15.4	85.6	525	6	BD147718
89	15.4	85.6	527	6	AR413963
90	15.4	85.6	527	6	AX970797
91	15.4	85.6	527	6	BD109516
92	15.4	85.6	608	6	AX079605

AC118955 Rattus no
X00876 Murine gene
AX435258 Sequence
AB038502 Caenorhab
AU297973 Mus muscu
AR052877 Sequence
AR127895 Sequence
AR288189 Sequence
AY044188 Mus muscu
AF161020 Mus muscu
AF051368 Mus muscu
M13873 Mouse p53 m
M13872 Mouse p53 m
AB017816 Mus muscu
M13874 Mouse p53 m
AB017815 Mus muscu
X00741 Mouse mRNA
AF151353 Mus muscu
AB021961 Mus muscu
AB020317 Mus muscu
AY212017 Mus muscu
X01237 Mouse mRNA
BC005448 Mus muscu
AC023805 Mus muscu
AL731687 Mouse DNA
AL445243 Human DNA
AC102741 Mus muscu
BX324147 Dantio rer
AC074146 Mus muscu
AC148732 Callithrix
AC074149 Mus muscu
AC105597 Rattus no
AC098078 Rattus no
J04357 Red clover
AC009656 Homo sapi
AC011750 Homo sapi
AL133292 Arabidops
AC019317 Homo sapi
AC027450 Homo sapi
AL592221 Human DNA
AY510475 Gallid he
AF147806 Gallid he
AF243438 Gallid he
AC015756 Homo sapi
AC061999 Homo sapi
AC131935 Homo sapi
AC119278 Mus muscu
AC137798 Homo sapi
AP326781 Tricicum
AC108811 Mus muscu
AC139741 Gallus ga
AF027558 Actus nan
AF196265 Rattus no
AF196264 Rattus no
AF196260 Rattus no
AF196262 Rattus no
AF196263 Rattus no
AF196261 Rattus no
AF196255 Rattus no
AF196259 Rattus no
X08525 R. norvegicu
X88527 R. norvegicu
A42904 Sequence 36
AR271438 Sequence
U07878 Mus musculu
U76836 Rattus norv
AX867656 Sequence
BD147718 Primer fo
AR413963 Sequence
AX970797 Sequence
BD109516 EST and e
AX079605 Sequence

C 93	15.4	85.6	826	8	BT004454	165	15.4	85.6	110000	2	AC140812_3	Continuation (4 of
94	15.4	85.6	833	8	HLI224591	167	15.4	85.6	110000	2	LMFLCR36_08	Continuation (9 of
95	15.4	85.6	870	6	CQ780681	168	15.4	85.6	110586	2	AC136235	AC136235 Rattus no
96	15.4	85.6	870	6	CQ782219	169	15.4	85.6	113254	2	AL137009	AL137009 Human DNA
97	15.4	85.6	870	6	BD125350	170	15.4	85.6	114302	2	AL162413	AL162413 Human DNA
98	15.4	85.6	870	6	BD126928	171	15.4	85.6	118017	2	AC147013	AC147013 Medicago
C 99	15.4	85.6	886	10	MMU04624	172	15.4	85.6	119865	10	AC0004407	AC0004407 Mus muscu
100	15.4	85.6	994	8	HSI224549	173	15.4	85.6	126470	2	AL1390234	AL1390234 Human DNA
101	15.4	85.6	1007	8	HPR224547	174	15.4	85.6	132597	2	AC074014	AC074014 Homo sapi
102	15.4	85.6	1018	8	HBA224553	175	15.4	85.6	132715	5	BX465189	BX465189 Zebrafish
103	15.4	85.6	1018	8	HLO224551	176	15.4	85.6	136292	2	AC144322	AC144322 Macaca mu
104	15.4	85.6	1018	8	HP1224556	177	15.4	85.6	136625	2	AC120597	AC120597 Homo sapi
105	15.4	85.6	1019	8	HHA224548	178	15.4	85.6	138940	8	AC063972	AC063972 Genomic S
106	15.4	85.6	1019	8	HSO224557	179	15.4	85.6	142256	2	AC079297	AC079297 Homo sapi
107	15.4	85.6	1033	8	HVE224561	180	15.4	85.6	144335	2	AC136445	AC136445 Homo sapi
108	15.4	85.6	1034	8	HBT224564	181	15.4	85.6	146304	2	AC061967	AC061967 Homo sapi
C 109	15.4	85.6	1069	8	BT002894	182	15.4	85.6	148807	2	AC007028	AC007028 Homo sapi
C 110	15.4	85.6	1078	8	AY084550	183	15.4	85.6	150152	2	AL355502	AL355502 Homo sapi
C 111	15.4	85.6	1296	10	RRTCRAR	184	15.4	85.6	151500	2	AC008355	AC008355 Drosophill
112	15.4	85.6	1364	6	AX876492	185	15.4	85.6	153426	10	AC121949	AC121949 Mus muscu
113	15.4	85.6	1364	6	BD156157	186	15.4	85.6	153572	5	BX323992	BX323992 Zebrafish
114	15.4	85.6	1364	6	AK027524	187	15.4	85.6	154034	10	AL772336	AL772336 Mouse DNA
115	15.4	85.6	1478	9	BC041640	188	15.4	85.6	154060	9	AL662885	AL662885 Human DNA
116	15.4	85.6	1490	6	CQ783924	189	15.4	85.6	157051	9	AC011095	AC011095 Homo sapi
117	15.4	85.6	1490	6	BD127903	190	15.4	85.6	159094	5	AL929193	AL929193 Zebrafish
118	15.4	85.6	1490	6	AK075317	191	15.4	85.6	159274	9	AC112651	AC112651 Homo sapi
119	15.4	85.6	1496	9	BC012140	192	15.4	85.6	159830	10	AC145267	AC145267 Mus muscu
120	15.4	85.6	1517	9	BC004911	193	15.4	85.6	161566	2	AC136462	AC136462 Rattus no
121	15.4	85.6	1535	6	BD063240	194	15.4	85.6	162874	2	AC101892	AC101892 Mus muscu
C 122	15.4	85.6	1598	11	BV176766	195	15.4	85.6	164459	2	AC105335	AC105335 Mus muscu
123	15.4	85.6	1664	6	AX597496	196	15.4	85.6	166312	9	AC067807	AC067807 Homo sapi
C 124	15.4	85.6	1664	6	AF052149	197	15.4	85.6	166988	10	AC144552	AC144552 Mus muscu
C 125	15.4	85.6	1670	5	BC094935	198	15.4	85.6	167830	2	AC138799	AC138799 Homo sapi
C 126	15.4	85.6	1755	5	BD192217	199	15.4	85.6	169260	2	AC027768	AC027768 Homo sapi
C 127	15.4	85.6	1848	5	BC055232	200	15.4	85.6	172616	2	AC132232	AC132232 Mus muscu
C 128	15.4	85.6	1884	3	AB060223	201	15.4	85.6	172914	2	AC129563	AC129563 Mus muscu
C 129	15.4	85.6	1975	3	AY060981	202	15.4	85.6	173069	2	AC141248	AC141248 Homo sapi
C 130	15.4	85.6	2003	3	BC009866	203	15.4	85.6	173217	2	AC027785	AC027785 Homo sapi
C 131	15.4	85.6	2045	3	BT001563	204	15.4	85.6	173382	2	AC145309	AC145309 Homo sapi
C 132	15.4	85.6	2214	3	COS98503	205	15.4	85.6	173675	2	AC090457	AC090457 Homo sapi
C 133	15.4	85.6	2243	3	AY625501	206	15.4	85.6	173913	2	AC139181	AC139181 Pan trogl
C 134	15.4	85.6	2243	3	COS92971	207	15.4	85.6	174092	2	AC069129	AC069129 Homo sapi
C 135	15.4	85.6	3336	6	BC048021	208	15.4	85.6	174930	2	AC145889	AC145889 Pan trogl
C 136	15.4	85.6	3703	6	COS72543	209	15.4	85.6	175161	2	BX927356	BX927356 Danio rer
C 137	15.4	85.6	4532	6	COS958502	210	15.4	85.6	175524	2	AC012108	AC012108 Homo sapi
C 138	15.4	85.6	5086	8	AY675586	211	15.4	85.6	177433	2	AP001857	AP001857 Homo sapi
C 139	15.4	85.6	6415	9	HGM806183	212	15.4	85.6	177930	2	AP001458	AP001458 Homo sapi
C 140	15.4	85.6	6446	9	COS92970	213	15.4	85.6	178438	2	AC149105	AC149105 Homo sapi
C 141	15.4	85.6	34867	8	AP002030	214	15.4	85.6	178900	2	AL139105	AL139105 Pepto anu
C 142	15.4	85.6	35353	2	AC020055	215	15.4	85.6	179532	2	AL135926	AL135926 Homo sapi
C 143	15.4	85.6	39575	2	AC145683	216	15.4	85.6	179711	2	AC102797	AC102797 Homo sapi
C 144	15.4	85.6	41131	2	AC091480_6	217	15.4	85.6	180242	2	AC044784	AC044784 Homo sapi
C 145	15.4	85.6	4558	9	AC105390	218	15.4	85.6	180937	2	AC141310	AC141310 Homo sapi
C 146	15.4	85.6	46643	10	AL173512	219	15.4	85.6	181142	3	AC0087035	AC0087035 Homo sapi
C 147	15.4	85.6	55520	9	HSBU488M4	220	15.4	85.6	181380	10	AC115123	AC115123 Mus muscu
C 148	15.4	85.6	56171	2	AC014872	221	15.4	85.6	182147	2	AC062032	AC062032 Homo sapi
C 149	15.4	85.6	61223	2	AC142438	222	15.4	85.6	182196	2	AP002773	AP002773 Homo sapi
C 150	15.4	85.6	65449	2	AC105402	223	15.4	85.6	182952	2	CNS06C80	AL331154 Human chr
C 151	15.4	85.6	69619	2	AC103638	224	15.4	85.6	183412	2	AP001899	AP001899 Homo sapi
C 152	15.4	85.6	81372	2	AC139056	225	15.4	85.6	184164	2	AC091020	AC091020 Homo sapi
C 153	15.4	85.6	81365	2	AC147867	226	15.4	85.6	185147	10	AC125521	AC125521 Mus muscu
C 154	15.4	85.6	87077	9	HUAC004097	227	15.4	85.6	18515	2	BX957362	BX957362 Danio rer
C 155	15.4	85.6	89934	8	ATAC009337	228	15.4	85.6	185670	3	AC008097	AC008097 Drosophill
C 156	15.4	85.6	93063	2	AP000644	229	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 157	15.4	85.6	93683	2	AC138685	230	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 158	15.4	85.6	100099	10	AE008684_3	231	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 159	15.4	85.6	108720	8	AB046436	232	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 160	15.4	85.6	109274	8	AC138010	233	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 161	15.4	85.6	110000	2	AC096087_2	234	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 162	15.4	85.6	110000	2	AC139568_3	235	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 163	15.4	85.6	110000	2	AC123419_2	236	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 164	15.4	85.6	110000	2	AC138332_2	237	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl
C 165	15.4	85.6	110000	2	AC140811_3	238	15.4	85.6	185683	2	BS000626	BS000626 Pan trogl

239	15.4	85.6	187415	2	AC026352	AC026352	Homo sapi	312	15.4	85.6	246199	2	AC127759	AC127759	Rattus no
240	15.4	85.6	187562	2	AC140452	AC140452	Mus muscu	313	15.4	85.6	246731	2	AC097211	AC097211	Rattus no
241	15.4	85.6	187741	2	AC087681	AC087681	Homo sapi	314	15.4	85.6	248926	2	AC111553	AC111553	Rattus no
242	15.4	85.6	188197	2	AC104892	AC104892	Mus muscu	315	15.4	85.6	249433	2	AC128523	AC128523	Rattus no
243	15.4	85.6	188406	9	AC140658	AC140658	Homo sapi	316	15.4	85.6	250147	10	AC122191	AC122191	Rattus no
244	15.4	85.6	188434	2	AC120068	AC120068	Homo sapi	317	15.4	85.6	250498	2	AC11813	AC11813	Rattus no
245	15.4	85.6	188734	2	AC069443	AC069443	Homo sapi	318	15.4	85.6	250810	2	AC103486	AC103486	Rattus no
246	15.4	85.6	189549	10	AC122880	AC122880	Mus muscu	319	15.4	85.6	251426	2	AC111816	AC111816	Rattus no
247	15.4	85.6	189689	10	AC123847	AC123847	Mus muscu	320	15.4	85.6	252435	2	AC094150	AC094150	Rattus no
248	15.4	85.6	190597	10	AC140371	AC140371	Mus muscu	321	15.4	85.6	252638	2	AC135932	AC135932	Rattus no
249	15.4	85.6	191533	10	AC123041	AC123041	Mus muscu	322	15.4	85.6	256038	2	AC013610	AC013610	Homo sapi
250	15.4	85.6	191854	2	AC141402	AC141402	Homo sapi	323	15.4	85.6	258648	2	AC130010	AC130010	Rattus no
251	15.4	85.6	192607	2	AC139984	AC139984	Rattus no	324	15.4	85.6	259730	2	AC094880	AC094880	Rattus no
252	15.4	85.6	193142	2	AC113861	AC113861	Rattus no	325	15.4	85.6	262239	2	AC147863	AC147863	Gallus ga
253	15.4	85.6	193291	10	AC132871	AC132871	Mus muscu	326	15.4	85.6	266232	2	AC094326	AC094326	Rattus no
254	15.4	85.6	193377	2	AC135187	AC135187	Mus muscu	327	15.4	85.6	274928	2	AC114717	AC114717	Rattus no
255	15.4	85.6	193513	2	AC074372	AC074372	Homo sapi	328	15.4	85.6	274993	2	AC133403	AC133403	Rattus no
256	15.4	85.6	196249	10	AC092202	AC092202	Mus muscu	329	15.4	85.6	276063	2	AC114458	AC114458	Rattus no
257	15.4	85.6	196838	2	AC115196	AC115196	Rattus no	330	15.4	85.6	281529	2	AC111407	AC111407	Rattus no
258	15.4	85.6	198348	9	AC141283	AC141283	Homo sapi	331	15.4	85.6	281797	3	AB003794	AB003794	Drosophill
259	15.4	85.6	198348	9	AC148025	AC148025	Homo sapi	332	15.4	85.6	282895	2	AC096056	AC096056	Rattus no
260	15.4	85.6	198992	10	AC124506	AC124506	Mus muscu	333	15.4	85.6	284938	2	AC104173	AC104173	Rattus no
261	15.4	85.6	199180	2	AC141078	AC141078	Homo sapi	334	15.4	85.6	285939	2	AC107599	AC107599	Rattus no
262	15.4	85.6	199868	9	AB019440	AB019440	Homo sapi	335	15.4	85.6	293854	2	AC126646	AC126646	Rattus no
263	15.4	85.6	200000	2	AC133765	AC133765	Rattus no	336	15.4	85.6	299420	2	AC094449	AC094449	Rattus no
264	15.4	85.6	200039	2	AC118255	AC118255	Mus muscu	337	15.4	85.6	299749	3	AB003606	AB003606	Drosophill
265	15.4	85.6	200859	2	AC124928	AC124928	Rattus no	338	15.4	85.6	299876	2	AC106196	AC106196	Rattus no
266	15.4	85.6	201214	9	AP002514	AP002514	Homo sapi	339	15.4	85.6	301660	2	AC112802	AC112802	Rattus no
267	15.4	85.6	201641	2	AC116494	AC116494	Mus muscu	340	15.4	85.6	304186	2	AC112438	AC112438	Rattus no
268	15.4	85.6	202375	2	AC090232	AC090232	Homo sapi	341	15.4	85.6	305730	2	AC114347	AC114347	Rattus no
269	15.4	85.6	202505	2	CR354554	CR354554	Danio rer	342	15.4	85.6	311823	2	AC095362	AC095362	Rattus no
270	15.4	85.6	202561	2	AC102839	AC102839	Mus muscu	343	15.4	85.6	313691	2	AC096152	AC096152	Rattus no
271	15.4	85.6	203260	2	AC011804	AC011804	Homo sapi	344	15.4	85.6	335400	2	AC131459	AC131459	Rattus no
272	15.4	85.6	205174	9	AC0110637	AC0110637	Rattus no	345	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
273	15.4	85.6	205515	2	AC146161	AC146161	Pan trogl	346	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
274	15.4	85.6	205885	9	AC141453	AC141453	Homo sapi	347	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
275	15.4	85.6	206056	2	AC145310	AC145310	Homo sapi	348	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
276	15.4	85.6	206536	2	AC133345	AC133345	Rattus no	349	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
277	15.4	85.6	206925	2	AC144632	AC144632	Mus muscu	350	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
278	15.4	85.6	207322	2	AC121502	AC121502	Mus muscu	351	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
279	15.4	85.6	207444	2	AC133405	AC133405	Rattus no	352	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
280	15.4	85.6	208406	2	AC125395	AC125395	Mus muscu	353	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
281	15.4	85.6	208938	10	AC114593	AC114593	Mus muscu	354	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
282	15.4	85.6	210843	2	AC140879	AC140879	Homo sapi	355	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
283	15.4	85.6	211018	2	AC090702	AC090702	Homo sapi	356	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
284	15.4	85.6	211250	2	AC118036	AC118036	Mus muscu	357	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
285	15.4	85.6	212345	2	AC118036	AC118036	Mus muscu	358	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
286	15.4	85.6	214428	2	AC114672	AC114672	Mus muscu	359	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
287	15.4	85.6	214504	2	AC141622	AC141622	Homo sapi	360	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
288	15.4	85.6	214511	2	AC129256	AC129256	Rattus no	361	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
289	15.4	85.6	215036	2	AC129359	AC129359	Rattus no	362	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
290	15.4	85.6	215213	10	AC111135	AC111135	Mus muscu	363	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
291	15.4	85.6	218856	2	AC122871	AC122871	Mus muscu	364	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
292	15.4	85.6	219412	10	AL772341	AL772341	Mus muscu	365	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
293	15.4	85.6	220173	9	AC012183	AC012183	Homo sapi	366	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
294	15.4	85.6	220595	2	AC128899	AC128899	Rattus no	367	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
295	15.4	85.6	224026	10	AC095402	AC095402	Rattus no	368	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
296	15.4	85.6	224221	10	AC117633	AC117633	Mus muscu	369	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
297	15.4	85.6	224890	2	AC118036	AC118036	Mus muscu	370	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
298	15.4	85.6	229336	2	AC131843	AC131843	Rattus no	371	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
299	15.4	85.6	229693	2	AC130032	AC130032	Rattus no	372	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
300	15.4	85.6	231171	2	AC132655	AC132655	Rattus no	373	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
301	15.4	85.6	232808	2	AC094043	AC094043	Rattus no	374	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
302	15.4	85.6	233447	2	AC103192	AC103192	Rattus no	375	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
303	15.4	85.6	235183	2	AC079427	AC079427	Mus muscu	376	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
304	15.4	85.6	235452	2	AC128416	AC128416	Rattus no	377	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
305	15.4	85.6	236164	10	AC101205	AC101205	Mus muscu	378	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
306	15.4	85.6	239282	2	AC096176	AC096176	Rattus no	379	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
307	15.4	85.6	239312	2	AC112572	AC112572	Rattus no	380	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
308	15.4	85.6	241231	2	AC137301	AC137301	Rattus no	381	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
309	15.4	85.6	243032	2	AC129167	AC129167	Rattus no	382	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
310	15.4	85.6	245698	2	AC124841	AC124841	Rattus no	383	15.4	85.6	336001	2	AC098369	AC098369	Rattus no
311	15.4	85.6	246021	10	AL663037	AL663037	Mus muscu	384	15.4	85.6	336001	2	AC098369	AC098369	Rattus no

385	15	83.3	182423	10	AC126942	C	458	14.8	82.2	2499	6	AX702458	AX702458	Sequence	
386	15	83.3	183101	9	AC021915	C	459	14.8	82.2	2498	6	AX275969	AX275969	Sequence	
387	15	83.3	186540	9	AC012478	C	460	14.8	82.2	2542	3	AY118550	AY118550	Drosophila	
388	15	83.3	186674	9	AC091544	C	461	14.8	82.2	2557	6	C0613581	C0613581	Sequence	
389	15	83.3	187234	2	AC114142	C	462	14.8	82.2	2832	6	C0577295	C0577295	Sequence	
390	15	83.3	187517	5	AY519500	C	463	14.8	82.2	2908	8	AY253257	AY253257	Diploclax	
391	15	83.3	187691	10	AC083816	C	464	14.8	82.2	4050	8	AF237210	AF237210	Drosophila	
392	15	83.3	196661	10	AC119967	C	465	14.8	82.2	4081	8	AK100406	AK100406	Oryza sat	
393	15	83.3	205350	10	AC078946	C	466	14.8	82.2	4081	8	AK100406	AK100406	Oryza sat	
394	15	83.3	207161	2	AC141175	C	467	14.8	82.2	4681	6	C0613580	C0613580	Sequence	
396	15	83.3	208600	9	AP002348	C	468	14.8	82.2	4725	6	C0613466	C0613466	Sequence	
397	15	83.3	212374	2	BX297371	C	469	14.8	82.2	5974	2	AC018252	AC018252	Drosophila	
398	15	83.3	215429	10	AC113542	C	470	14.8	82.2	9275	14	AB017037	AB017037	Himelob	
399	15	83.3	216449	2	AC051616	C	471	14.8	82.2	9275	14	AB183472	AB183472	Himelob	
400	15	83.3	220674	2	BX548000	C	472	14.8	82.2	11039	1	AB007268	AB007268	Sinorhizo	
401	15	83.3	221357	2	AP004246	C	473	14.8	82.2	11279	8	OSAS35073	AJ535073	Oryza sat	
402	15	83.3	221941	9	AC091060	C	474	14.8	82.2	12631	1	AE004252	AE004252	Vibrio ch	
403	15	83.3	222308	2	AC095280	C	475	14.8	82.2	14093	3	AE000860	AE000860	Methanoba	
404	15	83.3	224788	2	AP001905	C	476	14.8	82.2	18637	3	CBR225X01	AC084535	Caenorhab	
405	15	83.3	226505	2	AC006899	C	477	14.8	82.2	20809	3	CER04C1	AC015209	Caenorhab	
406	15	83.3	235904	2	AC102729	C	478	14.8	82.2	27365	2	AC015209	AC015209	Caenorhab	
407	15	83.3	237644	2	CR118653	C	479	14.8	82.2	30332	3	CER43C3	AC015209	Caenorhab	
408	15	83.3	240050	1	SA676684	C	480	14.8	82.2	30332	3	CER43C3	AC015209	Caenorhab	
409	15	83.3	241173	2	AC126150	C	481	14.8	82.2	38388	2	AC019820	AC019820	Caenorhab	
410	15	83.3	241781	2	AC122087	C	482	14.8	82.2	39216	3	CER0250	AC019820	Drosophila	
411	15	83.3	243619	2	AC132087	C	483	14.8	82.2	42035	6	C0363774	C0363774	Sequence	
412	15	83.3	259619	2	AC108308	C	484	14.8	82.2	42241	2	CER08B6	C0363774	Sequence	
413	15	83.3	266775	2	AC129793	C	485	14.8	82.2	47589	2	AC124053	AC124053	Mus muscu	
414	15	83.3	269788	10	AC125028	C	486	14.8	82.2	68663	2	AC116750	AC124053	Mus muscu	
415	15	83.3	272563	8	AC095469	C	487	14.8	82.2	69446	2	AC137105	AC137105	Mus muscu	
416	15	83.3	279448	8	AY661656	C	488	14.8	82.2	74149	3	AC123879_3	AC137105	Mus muscu	
417	15	83.3	289192	2	AC139506	C	489	14.8	82.2	85095	3	AC004573	AC004573	Drosophila	
418	15	83.3	297381	2	AE003540	C	490	14.8	82.2	89464	2	AC138543	AC138543	Magnapor	
419	15	83.3	349980	6	C0654010	C	491	14.8	82.2	90544	2	AC136262	AC136262	Rattus no	
15	83.3	349980	6	AX954527	Sequence	C	492	14.8	82.2	93899	2	AC105432	AC105432	Magnapor	
14.8	82.2	250	11	BTUAB228T	Sequence	C	493	14.8	82.2	95992	2	AC138324	AC138324	Mus sapi	
421	14.8	82.2	350	6	C0685634	Sequence	494	14.8	82.2	96330	8	AC14984	AC14984	Pomali b	
422	14.8	82.2	417	6	AX333593	Sequence	495	14.8	82.2	97862	5	BX36570	BX36570	Zebrafish	
423	14.8	82.2	417	6	AX334292	Sequence	C	496	14.8	82.2	100000	9	AP000071	AP000071	Hom sapi
424	14.8	82.2	417	6	AX407434	Sequence	C	497	14.8	82.2	102331	8	AC007120	AC007120	Arabidops
425	14.8	82.2	482	14	AY363364	Sequence	C	498	14.8	82.2	106073	9	AL136146	AL136146	Human DNA
426	14.8	82.2	527	5	CFCRCRS1	Sequence	499	14.8	82.2	110000	1	AB017283	AB017283	14	
427	14.8	82.2	635	11	BV026300	Sequence	499	14.8	82.2	110000	1	AC091229_04	AC091229_04	Continuation (5 of	
428	14.8	82.2	635	14	AY363365	Sequence	500	14.8	82.2	110000	2	AC091242_4	AC091242_4	Continuation (5 of	
429	14.8	82.2	687	14	HIV3B8LFR	Sequence									
430	14.8	82.2	714	14	AF239657	Sequence									
431	14.8	82.2	716	14	AF096643	Sequence									
432	14.8	82.2	763	14	AY221666	Sequence									
433	14.8	82.2	791	6	C0431966	Sequence									
434	14.8	82.2	824	14	AF462722	Sequence									
435	14.8	82.2	826	14	AF462721	Sequence									
436	14.8	82.2	928	6	C0729867	Sequence									
437	14.8	82.2	930	6	AF173378	Sequence									
438	14.8	82.2	966	6	AX275970	Sequence									
439	14.8	82.2	1002	6	CQ363925	Sequence									
440	14.8	82.2	1016	9	AY303790	Sequence									
441	14.8	82.2	1043	9	BC003013	Sequence									
442	14.8	82.2	1043	9	BC006504	Sequence									
443	14.8	82.2	1170	6	AX883719	Sequence									
444	14.8	82.2	1170	6	BD160491	Sequence									
445	14.8	82.2	1170	6	BD160491	Sequence									
446	14.8	82.2	1185	6	AX876719	Sequence									
447	14.8	82.2	1185	6	BD156279	Sequence									
448	14.8	82.2	1185	6	AK027569	Sequence									
449	14.8	82.2	1188	6	AX327430	Sequence									
450	14.8	82.2	1447	10	RATP53T02	Sequence									
451	14.8	82.2	1508	8	AF140552	Sequence									
452	14.8	82.2	2040	4	OCRNAP53	Sequence									
453	14.8	82.2	2137	8	AK064933	Sequence									
454	14.8	82.2	2340	3	AY437499	Sequence									
455	14.8	82.2	2368	3	AY634227	Sequence									
456	14.8	82.2	2388	3	C0604446	Sequence									
457	14.8	82.2	2488	3	DM047619	Sequence									

AC126942	Mus muscu	AC021915	Hom sapi	AC012478	Hom sapi	AC091544	Hom sapi	AC114142	Rattus no	AY519500	Gallus ga	AC083816	Mus muscu	AC119967	Mus muscu	AC078946	Mus muscu	AC141175	Rattus no	AP002348	Hom sapi	BX297371	Danio rer	AC113542	Mus muscu	AC051616	Mus muscu	BX548000	Danio rer	AP004246	Hom sapi	AC091060	Hom sapi	AC095280	Rattus no	AP001905	Hom sapi	AC006899	Caenorhab	AC102729	Mus muscu	CR118653	Danio rer	SA676684	Streptoco	AC126150	Rattus no	AC122087	Rattus no	AC132087	Rattus no	AC108308	Rattus no	AC129793	Rattus no	AC125028	Mus muscu	AC095469	Rattus no	AY661656	Sorghum b	AC139506	Hom sapi	AE003540	Drosophila	C0654010	Sequence	AX954527	Sequence	X98436	B. taurus an	C0685634	Sequence	AX333593	Sequence	AX334292	Sequence	AX407434	Sequence	AY363364	HIV-1 93S	CFCRCRS1	Sequence	X81458	C. fusus n	BV026300	S212P6009	AY363365	HIV-1 99S	L28880	Human	AF239657	HIV-1 iso	AF096643	HIV-1 str	AY221666	HIV-1 iso	C0431966	Sequence	AF462722	HIV-1 iso	AF462721	HIV-1 iso	C0729867	Sequence	AF173378	Hom sapi	AX275970	Sequence	CQ363925	Sequence	AY303790	Hom sapi	BC003013	Hom sapi	BC006504	Hom sapi	AX883719	Sequence	BD160491	Primter fo	AK024227	Hom sapi	AX876719	Sequence	BD156279	Primter fo	AK027569	Hom sapi	AX327430	Sequence	L12046	Rattus novy	AF140552	Lycopetera	X90592	O. cuniculus	AK064933	Oryza sat	AY437499	Typanosoma	AY634227	Oikopleur	C0604446	Sequence	DM047619	Drosophila
AX702458	Sequence	AX275969	Sequence	AY118550	Drosophila	C0613581	Sequence	C0577295	Sequence	AY253257	Diploclax	AF237210	Drosophila	AK100406	Oryza sat	C0604445	Sequence	C0613580	Sequence	C0613466	Sequence	AC018252	Drosophila	AB017037	Himelob	AB183472	Himelob	AB007268	Sinorhizo	AJ535073	Oryza sat	AE004252	Vibrio ch	AE000860	Methanoba	AC084535	Caenorhab	Z70267	Caenorhab	AC015209	Drosophila	Z47067	Caenorhab	AC019820	Drosophila	Z81453	Caenorhab	C0363774	Sequence	Z98854	Caenorhab	AC124053	Mus muscu	AC116750	Mus muscu	AC137105	Mus muscu	AC123879_3	Continuation (4 of	AC004573	Drosophila	AC138543	Magnapor	AC136262	Rattus no	AC105432	Magnapor	AC138324	Mus sapi	AC14984	Pomali b	BX36570	Zebrafish	AP000071	Hom sapi	AC007120	Arabidops	AL136146	Human DNA	AB017283	14	AC091229_04	Continuation (5 of	AC091242_4	Continuation (5 of																																																														

RESULT 1	RATP53T02/c	199 bp	DNA	linear	ROD 14-JUL-1993
LOCUS	RATP53T02				
DEFINITION	Rattus norvegicus tumor suppressor (p53) gene, exon 2.				
ACCESSION	L07904				
VERSION	L07904.1				
KEYWORDS	tumor suppressor.				
SEGMENT	2 of 8				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus				
REFERENCE	1 (bases 1 to 199)				
AUTHORS	Hulla J.E. and Schneider, R.P.				
TITLE	Structure of the rat p53 tumor suppressor gene				
JOURNAL	Nucleic Acids Res. 21 (3), 713-717 (1993)				
MEDLINE	93181268				
COMMENT	Original source text: Rattus norvegicus (strain Sprague-Dawley) DNA.				
FEATURES	From EMBL. entry VCRP001, dated 23-OCT-1992.				
source	Location/Qualifiers				
	1..199				
	/organism="Rattus norvegicus"				

PUBMED 9278438
 REFERENCE 2 (bases 1 to 1176)
 AUTHORS Mathupala,S.P.
 TITLE Direct Submission
 JOURNAL Submitted (21-FEB-1997) Biological Chemistry, Johns Hopkins
 University School of Medicine, 725 N. Wolfe Street, Baltimore, MD
 21205, USA

FEATURES
 source Location/Qualifiers
 1..1176
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /cell_line="AS-30D"
 /cell_type="ascites hepatoma"
 1..1176
 /function="enhances transcription of type II hexokinase"

CDS
 1..1176
 /codon_start=1
 /product="mutant p53"
 /protein_id="AAB80959.1"
 /db_xref="GI:1938365"
 /translation="MEDSQSMSTELPLSOETPSCIMKLLPPDIIPTTAGSPNSWE
 DLPFPOVAELLEGPERALQYSABAQEPGTAPAPAPASATPMPSSSVPSCKTQ
 SNIFHLGFQSGTAKSVCTYISLNLFCQLAKCPVOLWVTSTPPGTRVPMAL
 YKKSQNMTEVVRCPNHERCSGDGLAPPHLIRVEGNPYAEYLDROTFRHSVVVY
 EPPVSGSDYTTIHYKVMCSGCMGNRRPILITLIGDSSGNLIGRDFEVRVAC
 GRDRTEENFRKKEHCPELPGSAKRALPTSTSSPOOKKPLDEYFTLKRGRE
 RFEMFREINALELKDARAARESGDSRAHSYPRTKKGQSTSHKMKMIVGPDSD"

ORIGIN

Query Match 100.0%; Score 18; DB 10; Length 1176;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
 Db 18 CGACTGTGAATCCTCCAT 1

RESULT 5
 AX401815/c 1627 bp DNA linear PAT 06-JUN-2002
 LOCUS AX401815
 DEFINITION Sequence 1491 from Patent WO0210453.
 ACCESSION AX401815
 VERSION AX401815.1 GI:21337995
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1
 AUTHORS Mendrick,D., Porter,M.W., Johnson,K.R., Castle,A.L. and
 Elashoff,M.R.
 TITLE Molecular toxicology modeling
 JOURNAL Patent: WO 0210453-A 1491 07-FEB-2002;
 Gene Logic, Inc. (US)
 FEATURES
 source Location/Qualifiers
 1..1627
 /organism="Rattus norvegicus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10116"
 /note="EMBL/Genbank Accession No. X13058"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1627;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
 Db 41 CGACTGTGAATCCTCCAT 24

RESULT 6
 AX827834/c 1627 bp DNA linear PAT 12-DEC-2003
 LOCUS AX827834
 DEFINITION Sequence 568 from Patent EP1344834.
 ACCESSION AX827834
 VERSION AX827834.1 GI:39838022
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1
 AUTHORS Boess,F., Suter-Dick,L. and Wolf,D.
 TITLE Methods for the toxicity prediction of a compound
 JOURNAL Patent: EP 1344834-A 568 17-SEP-2003;
 F. HOFFMANN-LA ROCHE AG (CH)

FEATURES
 source Location/Qualifiers
 1..1627
 /organism="Rattus norvegicus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10116"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1627;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
 Db 41 CGACTGTGAATCCTCCAT 24

RESULT 7
 RNP53/c 1627 bp mRNA linear ROD 12-SEP-1993
 LOCUS RNP53
 DEFINITION Rat mRNA for nuclear oncoprotein p53.
 ACCESSION X13058
 VERSION X13058.1 GI:56828
 KEYWORDS oncoprotein p53.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 1627)
 AUTHORS Soussi,T., Caron de Fromental,C., Breugnot,C. and May,E.
 TITLE Nucleotide sequence of a cDNA encoding the rat p53 nuclear
 JOURNAL oncoprotein
 MEDLINE Nucleic Acids Res. 16 (23), 11384 (1988)
 PUBMED 89083585
 REFERENCE 2 (bases 1 to 1627)
 AUTHORS Soussi,T.
 TITLE Direct Submission
 JOURNAL Submitted (26-SEP-1988) Soussi T., Universite Pierre et Marie
 Curie, Unite d'Oncologie Molculaire, IRSC - CNRS, BP 08 94802
 Villejuif, France
 data kindly reviewed (09-Feb-1989) by Soussi T.

COMMENT
 FEATURES
 source Location/Qualifiers
 1..1627
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /cell_line="Py T21"
 24..1199
 /note="unnamed protein product; nuclear protein p53 (AA 1
 -391)"
 /codon_start=1
 /protein_id="CAA31457.1"
 /db_xref="GI:56829"
 /db_xref="GOA:P10361"

CDS

misc_feature	4305..4312	/note="NotI cloning site"	
misc_feature	4315..6027	/note="encodes yeast ADE2 reporter gene"	
Query Match	100.0%;	Score 18; DB 12; Length 9414;	
Best Local Similarity	100.0%;	Fred. No. 20;	
Matches	18; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	CGACTGTGAATCCCTCCAT 18	
Db	4290	CGACTGTGAATCCCTCCAT 4273	
RESULT 9	AC119115/c		
LOCUS			
DEFINITION	AC119115	165316 bp DNA linear HTG 19-NOV-2002	
ACCESSION	AC119115		
VERSION	AC119115.4	GI:25100662	
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
REFERENCE	1	(bases 1 to 165316)	
AUTHORS	Muzny,D,Marie., Metzker,M,Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amthun,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Baiswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buahy,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A., Chacko,U., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Bayes,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Faller,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Gante,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Giller,K., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howell,S., Huily,S., Hume,J., Idebrild,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowic,C., Kraft,C.L., Lebow,H., Levay,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenzulewa,L., Loulsegged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmood,M., Malloy,K., Mangun,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Matinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Mijna,R., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemelehu,O., Okunnu,G., Olarinpuason,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfennigoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., L., Prazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Rauter,M., Richards,S., Riggs,F., Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savary,G., Scheier,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorcelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A., Swatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,		

Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 165316)
 Worley, K.C.
 Direct Submission
 Submitted (25-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 165316)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:23616728.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GUXG
 Center clone name: CH230-320N23
 ----- Summary Statistics
 Assembly program: Phrap; Version 0.990329
 Consensus quality: 155971 bases at least Q40
 Consensus quality: 157321 bases at least Q30
 Consensus quality: 158253 bases at least Q20
 Estimated insert size: 159662; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 159511: contig of 159511 bp in length
 * 159512 159611: gap of unknown length
 * 159612 160669: contig of 1058 bp in length
 * 160670 160769: gap of unknown length
 * 160770 161864: contig of 1095 bp in length
 * 161865 161964: gap of unknown length
 * 161965 163701: contig of 1737 bp in length
 * 163702 163801: gap of unknown length
 * 163802 165316: contig of 1515 bp in length.
 Location/Qualifiers
 1. 165316
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-320N23"

misc_feature 1. 2103
 /note="wgs_end_extension
 clone_end:17"
 8259. 9143
 /note="clone_boundary
 clone_end:17
 site:
 end_sequence:BZ171094"
 46341. 159014
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ171096"

misc_feature
 end_sequence:BZ171094"
 46341. 159014
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ171096"

ORIGIN
 Query Match 100.0%; Score 18; DB 2; Length 165316;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CGACTGGAATCCCTCAT 18
 Db 14541 CGACTGGAATCCCTCAT 14524

RESULT 10
 AC134317
 LOCUS
 DEFINITION
 Rattus norvegicus clone RP31-421B16 strain Brown Norway, WORKING
 DRAFT SEQUENCE, 12 ordered pieces.
 AC134317
 AC134317.3 GI:31442441
 VERSION
 KEYWORDS
 HTG: HTGS PHASE2; HTGS DRAFT.
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

1 (bases 1 to 192326)
 Antonellis, A., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carls, K.,
 Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,
 Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
 Hurle, B., Idol, J.R., Karlins, E., Kwong, P., Laric, P., Lee-Lin, S.-O.,
 Legaspi, R., Maduro, Q.L., Maduro, V.B., Marulles, E.H., Mastello, C.,
 Maskeri, B., McDowell, J., Pagnitigan, C., Pearson, R., Portnoy, M.E.,
 Prasad, A., Redix-Dugue, N., Schandier, K., Schneider, M.G., Shan, K.,
 Sison, C., Stancipop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,
 Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 192326)
 Green, E.D.
 Direct Submission
 Submitted (25-SEP-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gailthersburg, MD 20877, USA
 3 (bases 1 to 192326)
 Green, E.D.
 Direct Submission
 Submitted (06-JUN-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gailthersburg, MD 20877, USA
 On Jun 6, 2003 this sequence version replaced gi:2753661.

----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc_zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: dce
 Center clone name: 421B16

The sequence data in this record represents an 'enhanced'
 version of a Phase 2 submission. Specifically, the indicated
 order and orientation of each sequence contig has been

established using one or more of the following: read-pair data from individual subclones; overlaps with neighboring clones; alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies; the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 190228 bases at least Q40
 Consensus quality: 190757 bases at least Q30
 Consensus quality: 191041 bases at least Q20
 Insert size: 182000; agarose-fp
 Insert size: 191226; sum-of-contigs
 Quality coverage: 12.13x in Q20 bases; agarose-fp
 Quality coverage: 11.55x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

```

1 10063: contig of 10063 bp in length
10064 10163: gap of unknown length
10164 12388: contig of 2225 bp in length
12389 12488: gap of unknown length
12489 29126: contig of 16638 bp in length
29127 29227: gap of unknown length
29227 31869: contig of 2643 bp in length
31870 31970: gap of unknown length
31970 45123: contig of 13154 bp in length
45124 45224: gap of unknown length
45224 54670: contig of 9447 bp in length
54671 54771: gap of unknown length
54771 62180: contig of 7410 bp in length
62181 62281: gap of unknown length
62281 66120: contig of 3840 bp in length
66121 66221: gap of unknown length
66221 114936: contig of 48716 bp in length
114937 115036: gap of unknown length
115037 168571: contig of 53535 bp in length
168572 168671: gap of unknown length
168671 189677: contig of 21206 bp in length
189678 189977: gap of unknown length
189978 192326: contig of 2349 bp in length.

```

FEATURES

SOURCE

```

1. 192326
   /organism="Rattus norvegicus"
   /mol_type="genomic DNA"
   /strain="Brown Norway"
   /db_xref="taxon:10116"
   /clone="RP31-421B16"
   /clone_1fb="RP31"

1. 13680
   /note="Clone overlaps with GenBank Accession Number AC136195 clone RP31-258K6 (center project name dcf)"

1. 10063
   /note="assembly_fragment"
   clone_end:sp6
   vector_side:left"

10164..12388
   /note="assembly_fragment"

12489..29126
   /note="assembly_fragment"

29227..31869
   /note="assembly_fragment"

```

```

misc_feature
31970..45123
/note="assembly_fragment"

misc_feature
45224..54670
/note="assembly_fragment"

misc_feature
54771..62180
/note="assembly_fragment"

misc_feature
62281..66120
/note="assembly_fragment"

misc_feature
66221..114936
/note="assembly_fragment"

misc_feature
115037..168571
/note="assembly_fragment"

misc_feature
168672..189677
/note="assembly_fragment"

misc_feature
189678..192326
/note="assembly_fragment"
clone_end:T7
vector_side:right"

```

ORIGIN

```

Query Match      100.0%; Score 18; DB 2; Length 192326;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACGTGATCCTCCAT 18
Db 50943 CGACGTGATCCTCCAT 50960

```

```

RESULT 11
AC136563/C
LOCUS
DEFINITION
AC136563
AC136563.3 GI:30521905
VERSION
KEYWORDS
HTG; HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 225077)
REFERENCE
Murny,D.Marie, Metzker,M.Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amun,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Bernhamed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Dayila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Derramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Georgievski,E., Geer,K., Gill,R., Garcia,M., Guerra,W., Guevara,M.,
Guanarone,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Huliyil,S., Hume,J., Idubaid,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kovacs,C., Kraft,C.L., Lebow,H., Levan,Z., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenzhuwa,L., Louissege,H., Lorado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,
Milosavljevic,A., Miner,G., Minja,B., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L.,

```

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Pasarek, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plöpper, F., Poldinger, A., Popovic, D., Prims, E., Pu, J.-L., Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Reijer, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ritz, S. J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soza, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Wortley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 225077)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (06-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 225077)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:25186706.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GHDQ
Center clone name: CH230-46E21
----- Summary Statistics
Assembly program: Atlas 3.0
Consensus quality: 214599 bases at least Q40
Consensus quality: 216978 bases at least Q30
Consensus quality: 218480 bases at least Q20
Estimated insert size: 225046; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a working draft sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
misc_feature
1. 1402
/note="wgs_contig"

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 18; DB 2; Length 225077;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTGTAATCTTCAT 18
|||||
Db 93482 CGACTGTAATCTTCAT 93465

RESULT 12
AL713860
LOCUS
DEFINITION
Mouse DNA sequence from clone RP23-446K18 on chromosome 11,
complete sequence.
AL713860
VERSION
AL713860.12 GI:27899576
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.
1 (bases 1 to 129489)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (24-JAN-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Jan 24, 2003 this sequence version replaced gi:21217846.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; SW, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep

Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
RP23-446K18 is from the RP21-23 Mouse BAC library constructed by the group of Pieter de Jong.
For further details see <http://www.choxi.org/bacpac/home.htm>
VECTOR: pBAc3.6
Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.
Location/Qualifiers

FEATURES

```

SOURCE
1. 129489
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-446X18"
/clone_1fb="RPCT-23"

ORIGIN
Query Match      94.4%; Score 17; DB 10; Length 129489;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2 GACTGTGAAATCTTCAT 18
Db      118503 GACTGTGAAATCTTCAT 118519

RESULT 13
AC119532/
LOCUS
DEFINITION
Rattus norvegicus clone CH230-268L16. *** SEQUENCING IN PROGRESS ***
AC119532
AC119532.5 GI:25009080
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
1 (bases 1 to 150816)
Munzy,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Bunay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesari,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Dayla,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Diaper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frieser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Healand,W., Hamli,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Huik,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Kapatthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kovacs,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensheva,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,B.,
Mawhinny,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,
Milosavljevic,A., Miner,G., Minja,B., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Natr,L.,
Nankervis,C., Neal,D., Newton,G., Olarnpunyaen,A., Pal,S., Parks,K.,
Nwokenemeh,O., Okwomu,G., Olarnpunyaen,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Prims,R., Pu,L.-L.,
Purao,M., Quiroz,J., Rachlin,B., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sander,M., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITL
JOURNAL
COMMENT
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23809253.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented. The sequence
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GVJF
Center clone name: CH230-268L16
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 140280 bases at least Q40
Consensus quality: 142072 bases at least Q30
Consensus quality: 143082 bases at least Q20
Estimated insert size: 144796; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Bactinated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_diffic_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 150816: contig of 150816 bp in length.
location/Qualifiers
1. 150816
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-268L16"
1. 1229
/notes="wgs end extension
clone end:5p6"
complement(4205..5074)
FEATURES
source
misc_feature
misc_feature
misc_feature

```

```

/note="clone boundary
clone_end:Sp6
site:
end_sequence:BT174726"
misc_feature
/clone="clone_boundary
clone_end:T7
site:
end_sequence:BT174725"
misc_feature
/clone="wgs_end_extension
clone_end:T7"

ORIGIN
Query Match          94.4%; Score 17; DB 2; Length 150816;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
DB 81343 GACTGTGAATCCTCCAT 81327

RESULT 14
CNS07YP2 181075 bp DNA linear ROD 23-MAR-2002
LOCUS Mus musculus chromosome 11 region in the Cm locus area
DEFINITION (D1Mit37-Scy46) clone 330D23 of library Caltech CITB-BAC from
chromosome 11 of Mus musculus (mouse).
ACCESSION AL713885
VERSION AL713885.1 GI:19701068
KEYWORDS 425A1-SP6; DDK syndrome; Ovum mutant; Scy41.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 181075)
AUTHORS Genoscope.
TITLES Direct Submission
JOURNAL Submitted (20-MAR-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
Cohen-Tannoudji M., Vandormael-Pournin S., Lebras S., Commailleau
F., Babinet C., Baldacci, P.
Unité de Biologie du Développement, CNRS URA 1960, Institut
Pasteur, 25 rue du Dr Roux, 75724 Paris cedex 15, France.
location/Qualifiers
1. 181075
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="330D23"
/clone_id="Caltech CITB-BAC"

ORIGIN
Query Match          94.4%; Score 17; DB 10; Length 181075;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCCTCCAT 18
DB 57477 GACTGTGAATCCTCCAT 57493

RESULT 15
AC094940 229363 bp DNA linear HTG 09-MAY-2003
LOCUS Rattus norvegicus clone CH230-6E9, *** SEQUENCING IN PROGRESS ***
DEFINITION 5 unordered pieces.
ACCESSION AC094940

```

```

VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 229363)
REFERENCE
Muzny,D.,Wente,M.,Lee,A.,Abramson,S.,Adams,C.,Alder,J.,
Allen,C.,Allen,H.,Albrooks,S.,Amit,A.,Anguiano,D.,
Anyalebechi,V.,Aoyagi,A.,Ayodeji,M.,Bac,E.,Bader,H.,
Baldwin,D.,Bandaranaik,D.,Barber,M.,Barnstead,M.,Benahmed,F.,
Biswal,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,
Bryant,N.,Buhay,C.,Burch,P.,Burrell,K.,Calderon,E.,
Cardenas,V.,Carter,K.,Cavazos,I.,Cesari,H.,Center,A.,
Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,
Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,
Davila,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dederich,D.,
Delgado,O.,Denson,S.,Derramo,C.,Ding,Y.,Dinh,H.,Divya,K.,
Draper,H.,Dugan-Rocha,S.,Dunn,A.,Dubin,K.,Dval,B.,Eaves,K.,
Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,
Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,
Fraser,C.M.,Gabisi,A.,Ganta,R.,Garcia,A.,Garnier,T.,Garza,M.,
Gebregorgis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,M.,Guevara,W.,
Gunnarane,P.,Healand,W.,Hamill,C.,Hamilton,C.,Hamilton,K.,
Harvey,Y.,Havlik,P.,Hawes,A.,Henderson,N.,Hernandez,J.,
Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,
Hollins,B.,Howells,S.,Huiy,S.,Hume,J.,Idlebird,D.,Jackson,A.,
Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jollivet,A.,
Karpachy,S.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Kovar,C.,
Kowis,C.,Kraft,C.L.,Lebow,H.,Levan,J.,Lewis,L.,Li,Z.,Liu,J.,
Liu,J.,Liu,W.,Liu,Y.,London,P.,Longacre,S.,Lopez,J.,
Lorenshewa,L.,Louiased,H.,Lozard,R.J.,Lu,X.,Ma,J.,
Maneswar,M.,Mahindartine,M.,Mahmoud,M.,Malloy,K.,Mangum,A.,
Mangum,B.,Mapua,P.,Martin,K.,Martin,R.,Martinez,E.,
Mawhinney,S.,McLeod,M.P.,McNeill,T.Z.,Meenen,E.,
Milosavljevic,A.,Miner,G.,Mintja,E.,Montemeyor,J.,Moore,S.,
Morgan,M.,Morris,K.,Morris,S.,Muniesa,M.,Murphy,M.,Nair,L.,
Nankervis,C.,Neal,D.,Newton,N.,Nguyen,N.,Norris,S.,
Nwaokelimen,O.,Okunolu,G.,Olarunpasegun,A.,Pal,S.,Parks,K.,
Pasternak,S.,Paul,H.,Perez,A.,Perez,L.,Pflamkoch,C.,
Plattner,F.,Polindexter,A.,Popovic,D.,Primus,E.,Pu,L.,
Plazom,M.,Quiroz,J.,Rachlin,E.,Reeves,K.,Regier,M.A.,Reigh,R.,
Rally,B.,Rally,M.,Ren,Y.,Reuter,M.,Richard,S.,Riggs,F.,
Rives,C.,Rodkey,T.,Rojas,A.,Rose,M.,Rose,R.,Ruiz,S.J.,
Sanders,M.,Savery,G.,Scherer,S.,Scott,G.,Shatman,S.,Shen,H.,
Shetty,J.,Shvartsbeyn,A.,Sisson,I.,Sitter,C.D.,Smay,D.,
Sneed,A.,Sodergren,E.,Song,X.-Z.,Sorelle,R.,Soes,J.,
Steinle,M.,Strong,R.,Sutton,A.,Swack,A.,Taber,P.,Taylor,C.,
Taylor,T.,Thomas,N.,Thomas,S.,Tingey,A.,Trejos,Z.,Umani,K.,
Valae,R.,Vera,V.,Villaana,D.,Waldron,L.,Walker,B.,Wang,J.,
Wang,Q.,Wang,S.,Warren,J.,Warren,K.,Wei,K.,White,F.,
Williams,G.,Willison,R.,Wleczky,R.,Wooden,H.,Worley,K.,
Wright,D.,Wright,R.,Wu,J.,Yakub,S.,Yen,J.,Yoon,L.,Yoon,Y.,
Yu,F.,Zhang,J.,Zhou,J.,Zhou,X.,Zhou,S.,Zhou,D.,von
Niederhausern,A.,Weiss,R.,Smith,D.R.,Holt,R.A.,Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 229363)
REFERENCE
Worley,K.C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 229363)
REFERENCE
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:22771362.
The sequence in this assembly is a combination of BAC based reads

```

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GBTU

Center clone name: CH230-6E9

----- Summary Statistics

Assembly program: Atlas;

Consensus quality: 208118 bases at least Q40

Consensus quality: 212027 bases at least Q30

Consensus quality: 214790 bases at least Q20

Estimated insert size: 226740; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html>).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 131104: contig of 131104 bp in length
* 131105 131204: gap of unknown length
* 131205 223436: contig of 92232 bp in length
* 223437 223536: gap of unknown length
* 223537 224657: contig of 1121 bp in length
* 224658 224757: gap of unknown length
* 224758 227566: contig of 2809 bp in length
* 227567 227666: gap of unknown length
* 227667 229363: contig of 1697 bp in length.

```

FEATURES

source

```

1. 229363
   /organism="Rattus norvegicus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10116"
   /clone="CH230-6E9"

```

ORIGIN

Query Match 94.4%; Score 17; DB 2; Length 229363;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GACTGTGATCTCCAT 18
 |||||
 Db 223332 GACTGTGATCTCCAT 223348

RESULT 16

AC094186/c

LOCUS AC094186 255706 bp DNA linear HTG 09-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-2K14, *** SEQUENCING IN PROGRESS ***,
 2 unordered pieces.

ACCESSION

AC094186 GI:30467744

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus

REFERENCE

AUTHORS

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 255706)

Murphy, D., Marie, Mezker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooke, S., Amin, A., Angiano, D., Anyalbechi, V., Ayagi, A., Ayodeji, M., Bacc, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biwaio, K., Blair, D., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, B., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., L., Davis, C., Day-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Dudin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Joilivet, A., Karpachy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, D., Lewis, L., Li, Z., Liu, J., Liu, Y., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshwaha, L., Lolliseed, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, B., Mangum, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, S., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, N., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoilemehe, O., Okunou, G., Olarunsoogun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poldexter, A., Popovic, D., Primus, B., Pu, L., L., Puzo, M., Quiroz, J., Rachin, B., Reeves, K., Regier, M., A., Reigh, R., Kelly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savory, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Slason, I., Sitter, C. D., Snaj, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Swalek, A., Taber, J., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Welser, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GADZ

Center clone name: CH230-2K14

Summary Statistics

Assembly program: Atlas;

Consensus quality: 244860 bases at least Q40

Consensus quality: 247754 bases at least Q30

Consensus quality: 249340 bases at least Q20

Estimated insert size: 268126; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved

* 1 254326: contig of 254326 bp in length
 * 254327 254426: gap of unknown length
 * 254427 255706: contig of 1280 bp in length.

Location/Qualifiers

1..255706

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-2K14"

/38438..43555

/note="wgs contig"

43606..46302

/note="wgs contig"

ORIGIN

Query Match 94.4%; Score 17; DB 2; Length 255706;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18

Db 56167 GACTGTGAATCTCCAT 56151

RESULT 17

AC103024

LOCUS Rattus norvegicus clone CH230-199E4, WORKING DRAFT SEQUENCE, 3

DEFINITION unorderd pieces.

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

AC103024 265074 bp DNA linear HTG 13-MAY-2003

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F., Bialski, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burck, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cessar, H., Chen, Z., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Bean, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagge, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabist, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladik, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulys, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Louisedge, H., Lozano, R.J., Lu, X., Ma, U., Maheshwari, M., Mahindaratne, M., Mamoud, M., Mallory, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Morgan, M., Morris, K., Morris, S., Munday, M., Murphy, M., Naik, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwachukwu, O., Okunolu, G., Olumide, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Polndexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodley, T., Rojas, A., Rose, M., Rose, R., Ruiz, J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Sma, J., Sma, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Wiczek, R., Wood, H., Worley, K., Williams, G., Willson, R., Wiczek, R., Wood, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhou, X., Zhou, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GIY

Center clone name: CH230-19984

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 243702 bases at least Q40

Consensus quality: 245482 bases at least Q30

Estimated insert size: 255715; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Rattified insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)

* NOTE: This sequence may represent more than one clone.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 262317: contig of 262317 bp in length

* 262318 262417: gap of unknown length

* 262367 263866: contig of 1349 bp in length

* 263867 265074: gap of unknown length

* 265074: contig of 1208 bp in length.

* Location/Qualifiers

* 1. 265074

* /organism="Rattus norvegicus"

* /mol_type="genomic DNA"

* /db_xref="taxon:10116"

* /clone="CH230-19984"

* 1. 2674

* /note="wgs end extension

* clone end: T7"

* complement(5474..5908)

* /note="Clone boundary

* site:BCORI

* end sequence: BH354676"

* complement(255145..258153)

* /note="Clone boundary

* clone end: Sp6

* site:BCORI

* end sequence: BH354677"

* 257056..258127

* /note="wgs end extension

* clone end: Sp6"

* 258629..262317

* /note="wgs end extension

* clone end: Sp6"

ORIGIN

Query Match 94.4% Score 17; DB 2; Length 265074;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 CGACTGTGATCTCTCA 17

DB 250504 CGACTGTGATCTCTCA 250520

RESULT 18
AC112826/c
LOCUS AC112826 267779 bp DNA linear HTG 08-OCT-2002

DEFINITION Rattus norvegicus clone CH230-21P15, *** SEQUENCING IN PROGRESS

AC112826

AC112826.3 GI:22856778

HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 267779)

REFERENCE

AUTHORS

Munzy D, Marie, Metker, M, Lee, Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Albrooks, S., Amn, A., Anguiano, D.,

Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, W., Barnhead, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Detamo, C., Ding, Y., Dinh, H., Divya, K.,

Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Georgiev, B., Geer, K., Gill, R., Grady, K., Guerra, W., Guevara, M.,

Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,

Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowals, C., Kraf, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensheewa, L., Louisedge, H., Lozano, R.J., Lu, X., Ma, J.,

Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,

Mawlinay, S., McLeod, M.P., McKell, T.Z., Meenan, B.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Natr, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwokelemeh, O., Okumu, G., Olarunsagoun, A., Pal, S., Parks, K.,

Paternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,

Plopper, F., Poindecker, A., Popovic, D., Primbs, B., Pu, L.,

Puazo, M., Quiroz, J., Rachlin, B., Reeves, K., Reiter, M.A., Reigh, R.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,

Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,

Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,

Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Snaj, D.,

Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J.,

Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, C., Taylor, C.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejo, Z., Umami, K.,

Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,

Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,

Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,

Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,

Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von

Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,

Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 267779)

Worley, K.C.

Direct Submission

Submitted (25-FEB-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 267779)

Direct Submission

Submitted (08-OCT-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

COMMENT

Baylor Plaza, Houston, TX 77030, USA

On Sep 14, 2002 this sequence version replaced gi:21738648.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRW
Center clone name: CH230-21P15

----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 223471 bases at least Q40
Consensus quality: 225356 bases at least Q30
Consensus quality: 226704 bases at least Q20
Estimated insert size: 241663; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 267779: contig of 267779 bp in length.
Location/Qualifiers

FEATURES
source
1..267779
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-21P15"
81136..83530
/note="wgs_contig"

misc_feature

ORIGIN

Query Match 94.4%; Score 17; DB 2; Length 267779;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GACTGTGAATCTCCAT 18
|||||

Db 90770 GACTGTGAATCTCCAT 90754

RESULT 19
AL627428 275838 bp DNA linear HTG 23-OCT-2002
LOCUS Mus musculus chromosome 11 clone RP23-386C14, 15 unordered pieces.
ACCESSION AL627428.23 GI:24366466
VERSION AL627428.23
KEYWORDS HTG; HTGS PHASE1; HTGS CANCELLED.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 275838)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS
TITLE
JOURNAL

COMMENT

Clark S.
Direct Submission
Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Oct 24, 2002 this sequence version replaced gi:20502228.

----- Genome Center

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bmk82J14

----- Summary Statistics

Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator; 1% of reads
Chemistry: Dye-terminator Big Dye; 98% of reads
Consensus quality: 270886 bases at least Q40
Consensus quality: 272427 bases at least Q40
Consensus quality: 273542 bases at least Q30
Insert size: 274438; sum-of-contigs
Insert size: 211998; 4.9% error; agarose-fp
Quality coverage: 10.60x in Q20 bases; sum-of-contigs quality
coverage: 13.87x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2049: contig of 2049 bp in length
2050 2149: gap of 100 bp
2150 7109: contig of 4960 bp in length
7110 7209: gap of 100 bp
7210 15665: contig of 8356 bp in length
15666 15665: gap of 100 bp
15666 27153: contig of 11488 bp in length
27154 27253: gap of 100 bp
27254 30924: contig of 3671 bp in length
30925 31024: gap of 100 bp
31025 37642: contig of 6618 bp in length
37643 37742: gap of 100 bp
37743 105267: contig of 67525 bp in length
105268 105367: gap of 100 bp
105368 251950: contig of 146583 bp in length
251951 252050: gap of 100 bp
252051 259747: contig of 7697 bp in length
259748 259847: gap of 100 bp
259848 262155: contig of 2308 bp in length
262156 262255: gap of 100 bp
262256 264502: contig of 2247 bp in length
264503 264602: gap of 100 bp
264603 266975: contig of 2373 bp in length
266976 267075: gap of 100 bp
267076 269446: contig of 2371 bp in length
269447 271564: gap of 100 bp
269547 271564: contig of 2018 bp in length
271565 271664: gap of 100 bp
271665 275838: contig of 4174 bp in length.

FEATURES

source

1..275838
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-386C14"
/clone_lib="RP23-386C14"
1..2049
/note="assembly_fragment:07407"

misc_feature


```

misc_feature      2150..7109
                    /note="assembly_fragment:07527
                    fragment_chain:1"
misc_feature      7210..15565
                    /note="assembly_fragment:07041
                    fragment_chain:1"
misc_feature      15666..27153
                    /note="assembly_fragment:07271
                    fragment_chain:1"
misc_feature      27254..30924
                    /note="assembly_fragment:07135
                    fragment_chain:2"
misc_feature      31025..37642
                    /note="assembly_fragment:06991
                    fragment_chain:2"
misc_feature      37743..105267
                    /note="assembly_fragment:03775"
                    /note="assembly_fragment:05020"
misc_feature      105368..251950
                    /note="assembly_fragment:05020"
misc_feature      252051..259747
                    /note="assembly_fragment:05467"
misc_feature      259848..262155
                    /note="assembly_fragment:05550"
misc_feature      262256..264502
                    /note="assembly_fragment:05948"
misc_feature      264603..266975
                    /note="assembly_fragment:06013"
misc_feature      267076..269446
                    /note="assembly_fragment:06040"
misc_feature      269547..271564
                    /note="assembly_fragment:07428"
misc_feature      271665..275838
                    /note="assembly_fragment:07590"

```

ORIGIN

```

Query Match      94.4% Score 17; DB 2; Length 275838;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2 GACTGTGATCCTCCAT 18
        |||||
Db      155872 GACTGTGATCCTCCAT 155888

```

```

RESULT 20
AC118955/AC118955/
LOCUS      AC118955
DEFINITION Rattus norvegicus clone CH230-21006, *** SEQUENCING IN PROGRESS
            *** 4 unordered pieces.
ACCESSION AC118955
VERSION   AC118955.6 GI:23269762
KEYWORDS  HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 328868)
REFERENCE 1
AUTHORS   Munuy D,Marie., Metzker M,lee., Abramson S., Adams C., Alder J.,
            Allen C., Allen H., Albrooks S., Amin A., Angilano D.,
            Anyalebech V., Aoyagi A., Ayodeji M., Baca B., Baden H.,
            Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
            Biewald K., Blair J., Blankenburg K., Blyth P., Brown M.,
            Bryant N., Buhay C., Burch P., Butrell K., Calderon E.,
            Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
            Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
            Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
            Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
            Delgado O., Denison S., Deramo C., Ding Y., Dinh H., Divya K.,
            Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
            Egan A., Esecotto M., Eugene C., Evans C.A., Falls T., Fan G.,
            Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,

```

```

TITLE      JOURNAL
REFERENCE 1
AUTHORS   Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
            Gebregregis B., Geer K., Gill R., Grady M., Guerra W., Guevara W.,
            Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,
            Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez D.,
            Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M.,
            Hollins B., Howells S., Hulik S., Hume J., Idlebird D., Jackson A.,
            Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolyet A.,
            Karpach S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
            Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
            Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
            Lorensuhewa L., Louised H., Lozada R.J., Lu X., Ma J.,
            Maheshwari M., Mahindartne M., Mahmood M., Mallory K., Mangum A.,
            Mangum B., Mapua P., Martin K., Martin R., Martinez B.,
            Mawhinney S., McLeod M.P., McNeill T.Z., Meenen B.,
            Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S.,
            Morgan M., Morris K., Morris S., Mundaasa M., Murphy M., Natl L.,
            Nankervil C., Neal D., Newton N., Nguyen N., Norris S.,
            Nockelmech O., Okwoum G., Olarnpungoon A., Pal S., Parks K.,
            Pasternak S., Paul H., Perez A., Perez L., Pfannkuch C.,
            Plopper F., Poindexter A., Popovic D., Prims E., Pu L.L.,
            Puzo M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R.,
            Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs F.,
            Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J.,
            Sanders M., Savery G., Scherer S., Scott G., Shatman S., Shen H.,
            Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajls D.,
            Sneed A., Sodergren E., Song X.-Z., Sorelle R., Sosa J.,
            Steimle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C.,
            Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Umanak K.,
            Valas R., Vera V., Villalana D., Waldron L., Walker B., Wang J.,
            Wang Q., Wang S., Warren J., Warren J., Wei X., White F.,
            Williams G., Willson R., Wiczysk R., Wooden H., Worley K.,
            Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V.,
            Yu F., Zhang J., Zhou J., Zhou X., Zhao S., Zhou D., von
            Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O.,
            Weinstein G., and Gibbs R.A.
            Direct Submission
            Unpublished
            2 (bases 1 to 328868)
            Worley K.C.
            Direct Submission
            Submitted (22-APR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            3 (bases 1 to 328868)
            Rat Genome Sequencing Consortium.
            Direct Submission
            Submitted (09-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            On Sep 23, 2002 this sequence version replaced gi:21746759.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads. Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.

```

```

COMMENT
REFERENCE 1
AUTHORS   Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
            Gebregregis B., Geer K., Gill R., Grady M., Guerra W., Guevara W.,
            Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,
            Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez D.,
            Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M.,
            Hollins B., Howells S., Hulik S., Hume J., Idlebird D., Jackson A.,
            Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolyet A.,
            Karpach S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
            Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
            Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
            Lorensuhewa L., Louised H., Lozada R.J., Lu X., Ma J.,
            Maheshwari M., Mahindartne M., Mahmood M., Mallory K., Mangum A.,
            Mangum B., Mapua P., Martin K., Martin R., Martinez B.,
            Mawhinney S., McLeod M.P., McNeill T.Z., Meenen B.,
            Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S.,
            Morgan M., Morris K., Morris S., Mundaasa M., Murphy M., Natl L.,
            Nankervil C., Neal D., Newton N., Nguyen N., Norris S.,
            Nockelmech O., Okwoum G., Olarnpungoon A., Pal S., Parks K.,
            Pasternak S., Paul H., Perez A., Perez L., Pfannkuch C.,
            Plopper F., Poindexter A., Popovic D., Prims E., Pu L.L.,
            Puzo M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R.,
            Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs F.,
            Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J.,
            Sanders M., Savery G., Scherer S., Scott G., Shatman S., Shen H.,
            Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajls D.,
            Sneed A., Sodergren E., Song X.-Z., Sorelle R., Sosa J.,
            Steimle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C.,
            Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Umanak K.,
            Valas R., Vera V., Villalana D., Waldron L., Walker B., Wang J.,
            Wang Q., Wang S., Warren J., Warren J., Wei X., White F.,
            Williams G., Willson R., Wiczysk R., Wooden H., Worley K.,
            Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V.,
            Yu F., Zhang J., Zhou J., Zhou X., Zhao S., Zhou D., von
            Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O.,
            Weinstein G., and Gibbs R.A.
            Direct Submission
            Unpublished
            2 (bases 1 to 328868)
            Worley K.C.
            Direct Submission
            Submitted (22-APR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            On Sep 23, 2002 this sequence version replaced gi:21746759.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads. Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            Project Information
            Center project name: CH230-21006
            Center clone name: CH230-21006
            Summary Statistics
            Assembly program: Phrap; version 0.990329
            Consensus quality: 218733 bases at least Q40
            Consensus quality: 218733 bases at least Q30
            Consensus quality: 218850 bases at least Q20
            Estimated insert size: 256170; sum-of-contigs estimation
            Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

```


ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea; Rhabditidae; Pelodermineae; Caenorhabditis.

REFERENCE 1 (sites)
AUTHORS Hirabayashi, Y., Hayama, K. and Kasai, K.
TITLE Novel galectins found in *C. elegans*
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 836)
AUTHORS Hirabayashi, Y., Hayama, K. and Kasai, K.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-2000) Jun Hirabayashi, Faculty of Pharmaceutical Sciences, Teikyo University, Department of Biological Chemistry; Suwarashi 1091-1, Sagami-cho, Kanagawa 199-0195, Japan (E-mail: j-hirabayashi@pharm.teikyo-u.ac.jp, Tel: 81-426-85-3741, Fax: 81-426-85-3742)

FEATURES
source Location/Qualifiers
1..836
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
1..836
/gene="lec-11"
1..22
/gene="lec-11"
/note="SL1 trans-spliced leader sequence"
97..795
/gene="lec-11"
/codon_start=1
/product="galectin LEC-11"
/protein_id="BAB11967.1"
/db_xref="GI:9857641"
/translation="MAYIENPPIPLTPIYEGKREHSSVIEHGEVFGPGGPIYER
PTKNGVALHISVNGSYGQNVIVFNHLDRGWHREHQNITMRERPCRIHNEHRK
YSIHVDGHHIGHHKKCPKRIVAITVAGDRLVGRIHFNPKHNGGTEIGVMQMPV
GPTIVVAQPPPMIVPQPIVPPVQIQIMQWQPIYSPVPIVPTATPIITYQPEV
IYMDGYHHRHHY"
822..827
/gene="lec-11"
polyA_signal

ORIGIN
Query Match 91.1%; Score 16.4; DB 3; Length 836;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGGAATCCTCCAT 18
|||||
Db 223 CGACTGGAATCCTCCCT 206

RESULT 24
MMU297973/c 1096 bp mRNA linear ROD 24-NOV-2000
LOCUS MMU297973
DEFINITION Mus musculus mRNA for transformation related protein 53 (Trp53 gene).
ACCESSION AJ297973
VERSION AJ297973.1 GI:11342598
KEYWORDS transformation related protein 53; Trp53 gene.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Rothbarth, K. and Werner, D.
TITLE Proteasome-mediated degradation antagonizes critical levels of the apoptosis-inducing CID protein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1096)
AUTHORS Werner, D.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2000) Werner D., Biochemistry of the Cell, German Cancer Research Center, Im Neuenheimer Feld 280, Heidelberg, 69120, GERMANY

FEATURES Location/Qualifiers

source 1..1096
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="EK398"
/tissue type="Ehrlich ascites tumor"
1..1096
/gene="p53"
1..924
/gene="p53"
/note="premature termination due to the deletion (compared to wildtype)"
/product="transformation related protein 53"
/codon_start=1
/protein_id="CAC17147.1"
/db_xref="GI:11342599"
/db_xref="GOA:Q9ER40"
/db_xref="TrEMBL:Q9ER40"
/translation="MSAMBSQSDISLEPLPSORTFSLWKLPPEDILSPHGMDDL
LLPDVBERPFGESBALRVSGAPAAQDPVTETPGPVAPATPMPPLSPFSQXTYOC
NYGHLAPLOSGLKSVCTYSPLNKLFCQIATCPVQLVSAITPFGSRVRLMALT
KSGQMTVEVARCPHHRCSDDGLAPQHLIRVEGNLYPEHLEDROTFRHSVVVPE
PPEAGSEYTIHYKYMCSNCSGCMNRPIITITLEDSPGNLIGRDSFEVRVACPG
RDRRTBEENFRKKEVLCPELPPGSAKRDPR"

ORIGIN
Query Match 91.1%; Score 16.4; DB 10; Length 1096;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGGAATCCTCCAT 18
|||||
Db 27 CGACTGGAATCCTCCAT 10

RESULT 25
AR052877/c 1173 bp DNA linear PAT 29-SEP-1999
LOCUS AR052877
DEFINITION Sequence 214 from patent US 5833975.
ACCESSION AR052877
VERSION AR052877.1 GI:5977739
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1173)
AUTHORS Paoletti, B., Tartaglia, J. and Cox, W. I.
TITLE Canaripox virus expressing cytokine and/or tumor-associated antigen DNA sequence
JOURNAL Patent: US 5833975-A 214 10-NOV-1998;
FEATURES Location/Qualifiers
source 1..1173
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 91.1%; Score 16.4; DB 6; Length 1173;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGGAATCCTCCAT 18
|||||
Db 27 CGACTGGAATCCTCCAT 10

RESULT 26
AR127895/c 1173 bp DNA linear PAT 16-MAY-2001
LOCUS AR127895
DEFINITION Sequence 45 from patent US 6183752.
ACCESSION AR127895
VERSION AR127895.1 GI:14115557
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1173)
AUTHORS Epstein,S.B., Finkel,T., Speir,E., Zhou,Y.Fu., Zhu,J., Erdile,L.
and Pincus,S.
TITLE Restenosis/atherosclerosis diagnosis, prophylaxis and therapy
JOURNAL Patent: US 6183752-A 45 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..1173
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 91.1%; Score 16.4; DB 6; Length 1173;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
Db 27 CGACTGTGACTCCTCCAT 10
|||||
|||||

RESULT 27
AR288189/c 1173 bp mRNA linear PAT 12-JUN-2003
LOCUS
DEFINITION Sequence 214 from patent US 6537594.
ACCESSION AR288189
VERSION AR288189.1 GI:31675468
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1173)
Paoletti,E., Tartaglia,J. and Cox,W.I.
TITLE Vaccina virus comprising cytokine and/or tumor associated antigen
JOURNAL Patent: US 6537594-A 214 25-MAR-2003;
FEATURES Location/Qualifiers
source 1..1173
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 91.1%; Score 16.4; DB 6; Length 1173;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
Db 27 CGACTGTGACTCCTCCAT 10
|||||
|||||

RESULT 28
AY044188/c 1176 bp mRNA linear ROD 21-DEC-2001
LOCUS
DEFINITION Mus musculus transformation related protein 53 (Trp53) mRNA,
complete cds.
ACCESSION AY044188
VERSION AY044188.1 GI:15375071
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 1176)
Irminger-Finger,I., Feki,A., Jefford,C.E., Soriano,D.V., Jacom,W., Montesano,R. and
Krause,K.-H.
TITLE Identification of BARD1 as Mediator between Proapoptotic Stress and
p53-Dependent Apoptosis
JOURNAL Mol. Cell 8 (6), 1255-1266 (2001)
PUBMED 11779501
REFERENCE 2 (bases 1 to 1176)

AUTHORS Irminger-Finger,I., Feki,A., Jefford,C.E. and Molnarfi,N.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2001) Geriatrics, Univ. of Geneva, 2 ch.
Petit-Bel-Air, Geneva CH-1225, Switzerland
FEATURES Location/Qualifiers
source 1..1176
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_line="FAC-2"
/cell_type="epithelial"
/tissue_type="mammary gland"
1..1176
/gene="Trp53"
1..1176
/gene="Trp53"
/note="p53; tumor suppressor"
/codon_start=1
/product="transformation related protein 53"
/protein_id="P4K94783.1"
/db_xref="GI:15375072"
/translation="MTAMERSQSDISLEPLSQETFSGLWKLPPEDILPSPHCMDDL
LPLDVEEPFEGSEALRVSGAPADOPVETGTPVAPAPATPMSFSPQKTYG
NYGPHLGILOSGLAKSYMCTVSPPLKLKFCGLAKTQVQLMWATPSPAGSRVAMATY
KKSQMTREIVRRCPHHRCSDDGLAPPHLITVEGNLYREYLEDQTPHSHVVDYE
PPRAGSYTHIHKICNCSQWGMRRPILITITLEDSSGNLGRDSFVRVACAPG
RDMRTREENRKKVLCPELPPSARALPTCTSAAPQPKKPLDGEVFTLKIKGRKR
FEMFREINELBIKDAHATBESGDSRAHSYLTQTKKGQSTSRHKTMVKKVGPDDSD"

ORIGIN
Query Match 91.1%; Score 16.4; DB 10; Length 1176;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCCTCCAT 18
Db 27 CGACTGTGACTCCTCCAT 10
|||||
|||||

RESULT 29
AF161020/c 1183 bp mRNA linear ROD 18-JUL-1999
LOCUS
DEFINITION Mus musculus tumor suppressor p53 (p53) mRNA, complete cds.
ACCESSION AF161020
VERSION AF161020.1 GI:5524685
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 1183)
Jiang,G.C., Yuan,L.Z., Wei,K. and Guo,X.M.
TITLE Sequence of tumor suppressor gene p53 in radiosensitive mouse breast
carcinoma cell line SX-9
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1183)
Jiang,G.C., Yuan,L.Z., Wei,K. and Guo,X.M.
AUTHORS Direct Submission
TITLE Submitted (21-JUN-1999) Institute of Radiation Medicine, Academy of
Military Medical Sciences, No.27 Taiping Road, Haiding District,
Beijing 100850, P. R. China
JOURNAL Location/Qualifiers
source 1..1183
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_line="SX-9"
/tissue_type="radiosensitive breast carcinoma"
1..1183
/gene="p53"
11..1183
/gene="p53"
/codon_start=1

```

/product="tumor suppressor p53"
/protein_id="A044340.1"
/db_xref="GI:5524686"
/translation="MTAMESSQSDISLEPLSQETPSGLWKLPPEDILSPHCMDLL
LPDVEEFEGPSBALRVSGAPADPVTETPGVAPAPATPWLSPFVSQKTYQG
NYGFLGLQSTAKSVWCYSPPLNKLFCQALATCPVQLVWSATPPAGSRVAMAIY
KKSOMTEVVRCPHRCNSDGDGLAPPOH.IRVGNLYPEYLEDROTFRHSVVVYVE
PPKASRTTTHYKTMCNSSCGMNRPIITITLDBSSNLLGRDSFEVRCACRG
RDRRTBEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLKIGRKR
FEMFREINLEALKDAAHTESGDSRAHSYLTKKQGSTSRHKKTWKKVGPDS"

```

ORIGIN

```

Query Match      91.1%; Score 16.4; DB 10; Length 1183;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy      1 CGACTGTGATCTCTCCAT 18
        |||||
Db      37 CGACTGTGATCTCTCCAT 20

```

```

RESULT 30      AF051368      1208 bp      mRNA      linear      ROD 16-MAR-1998
LOCUS          AF051368
DEFINITION     Mus musculus tumor suppressor p53 mRNA, complete cds.
ACCESSION      AF051368
VERSION        AF051368.1 GI:2961246
KEYWORDS
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
AUTHORS        Frostejo, L., Nilsson, J., Wandzioch, E. and Heby, O.
TITLE          Sequence of tumor suppressor p53 in mouse lymphoid leukemia cell
               line L1210
JOURNAL         Unpublished
REFERENCE       2 (bases 1 to 1208)
AUTHORS        Frostejo, L., Nilsson, J., Wandzioch, E. and Heby, O.
TITLE          Direct Submission
JOURNAL         Submitted (27-FEB-1998) Dept. Cellular and Developmental Biology,
               Umea University, Hus 1, Umea 901 87, Sweden

```

FEATURES

```

source
location/Qualifiers

```

```

1..1208
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_line="L1210"
/issue_type="lymphoid leukemia"
26..1198
/note="MLP53; mutant"
/codon_start=1
/product="tumor suppressor p53"
/protein_id="AAC05704.1"
/db_xref="GI:2961247"
/translation="MTAMESSQSDISLEPLSQETPSGLWKLPPEDILSPHCMDLL
LPDVEEFEGPSBALRVSGAPADPVTETPGVAPAPATPWLSPFVSQKTYQG
NYGFLGLQSTAKSVWCYSPPLNKLFCQALATCPVQLVWSATPPAGSRVAMAIY
KKSOMTEVVRCPHRCNSDGDGLAPPOH.IRVGNLYPEYLEDROTFRHSVVVYVE
PPKASRTTTHYKTMCNSSCGMNRPIITITLDBSSNLLGRDSFEVRCACRG
RDRRTBEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLKIGRKR
FEMFREINLEALKDAAHTESGDSRAHSYLTKKQGSTSRHKKTWKKVGPDS"
263..266
/note="encodes substitution of alanine for wildtype
valine"

```

CDS

```

misc_feature
/note="encodes substitution of alanine for wildtype
valine"

```

ORIGIN

```

Query Match      91.1%; Score 16.4; DB 10; Length 1208;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy      1 CGACTGTGATCTCTCCAT 18
        |||||

```

```

Db      52 CGACTGTGATCTCTCCAT 35

```

```

RESULT 31      M05P53B/c      1241 bp      mRNA      linear      ROD 27-APR-1993
LOCUS          M05P53B
DEFINITION     Mouse p53 mRNA, complete cds, clone p53-m11.
ACCESSION      M13873
VERSION        M13873.1 GI:200200
KEYWORDS
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
AUTHORS        Arai, N., Nomura, D., Yokota, K., Wolf, D., Brill, E., Shohat, O. and
               Rotter, V.
TITLE          Immunologically distinct p53 molecules generated by alternative
               splicing
JOURNAL         Mol. Cell. Biol. 6 (9), 3232-3239 (1986)
MEDLINS        87064640
PUBMED         3023970
COMMENT        Original source text: Mouse (BALB/c) Meth A library, cDNA to mRNA,
               clone p53-m11.

```

FEATURES

```

source
location/Qualifiers

```

```

1..1241
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/map="11"
/issue_1ib="Meth A"
1..1241
/gene="p53"
69..1241
/gene="p53"
/codon_start=1
/protein_id="AAA39882.1"
/db_xref="GI:200201"
/translation="MTAMESSQSDISLEPLSQETPSGLWKLPPEDILSPHCMDLL
LPDVEEFEGPSBALRVSGAPADPVTETPGVAPAPATPWLSPFVSQKTYQG
NYGFLGLQSTAKSVWCYSPPLNKLFCQALATCPVQLVWSATPPAGSRVAMAIY
KKSOMTEVVRCPHRCNSDGDGLAPPOH.IRVGNLYPEYLEDROTFRHSVVVYVE
PPKASRTTTHYKTMCNSSCGMNRPIITITLDBSSNLLGRDSFEVRCACRG
RDRRTBEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLKIGRKR
FEMFREINLEALKDAAHTESGDSRAHSYLTKKQGSTSRHKKTWKKVGPDS"

```

ORIGIN

```

Query Match      91.1%; Score 16.4; DB 10; Length 1241;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy      1 CGACTGTGATCTCTCCAT 18
        |||||
Db      95 CGACTGTGATCTCTCCAT 78

```

```

RESULT 32      M05P53A/c      1285 bp      mRNA      linear      ROD 27-APR-1993
LOCUS          M05P53A
DEFINITION     Mouse p53 mRNA, complete cds, clone pc053.
ACCESSION      M13872
VERSION        M13872.1 GI:200198
KEYWORDS
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
AUTHORS        Arai, N., Nomura, D., Yokota, K., Wolf, D., Brill, E., Shohat, O. and
               Rotter, V.
TITLE          Immunologically distinct p53 molecules generated by alternative

```

JOURNAL Mol. Cell. Biol. 6 (9), 3232-3239 (1986)
MEDLINE 87064640
PubMed 3023970

COMMENT Original source text: Mouse (BALB/c) nontransformed helper T-cell,
CDNA to mRNA, clone pCD-p53.
Location/Qualifiers

FEATURES

source
1..1285
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/cell_type="nontransformed helper T-cell"
113..1285
/codon_start=1
/product="p53 protein"
/protein_id="AAA39881.1"
/db_xref="GI:200199"
/translation="MTAMESOSDISLEPLSGEFTSGMLKLLPREDILPSPHCMIDL
LLPDVFEFEGPSEALRVSGAPAAQDPVTETPPVAPAPATPPLSPVPSQKTYG
NYGHLGFLSGTAKSVCTYSPPLNKLFCQIAKTCVQLVWSATPPAGSRVAMALY
KKSQHTMEVVRRCPEHRCSDGDLAPPHLIRVEGNLYPEYLEDROTFRHSVVVPE
PPEAGSEYTIHYKYMCSNCSMGGMNRPIITITLEDSSGNILGRSPFVAVCACPG
RDRRTTEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLKRGKR
FEMFRLEINALELKDANATEBSGDSRAHSYLTKTKKGQSTSRHKKTVMKKVGPDS"

CDS

1 CGACTGGAATCCTCCAT 18
|||||
139 CGACTGACTCTCCAT 122

ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1285;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGGAATCCTCCAT 18
|||||
Db 139 CGACTGACTCTCCAT 122

RESULT 33
AB017816/c 1305 bp mRNA linear ROD 09-JUL-1999
LOCUS Mus musculus mRNA for p53, complete cds.
DEFINITION AB017816
AB017816.1 GI:5421812
p53.
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
AUTHORS Araki, R., Fukumura, R., Fujimori, A., Tatsumi, K. and Abe, M.
TITLE Characterization of DNA-PKcs null mutant 5X9
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1305)
Fujimori, A. and Abe, M.
DIRECT SUBMISSION
SUBMITTED (22-SEP-1998) Masumi Abe, National Institute of
Radiological Sciences, Dept. of Biology and Oncology, Anagawa
4-9-1, Inage-ku, Chiba, Chiba 263-8555, Japan
(E-mail: abemasumi@nirs.go.jp, Tel:043-206-3219,
Fax:043-251-4593)
Location/Qualifiers

FEATURES

source

1..1305
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_line="5X9"
/cell_type="epithelial"
/tissue_type="mammary carcinoma"
102..1374
/codon_start=1
/product="p53"
/protein_id="BAA82340.1"
/db_xref="GI:5421813"

CDS

/translation="MTAMESOSDISLEPLSGEFTSGMLKLLPREDILPSPHCMIDL
LLPDVFEFEGPSEALRVSGAPAAQDPVTETPPVAPAPATPPLSPVPSQKTYG
NYGHLGFLSGTAKSVCTYSPPLNKLFCQIAKTCVQLVWSATPPAGSRVAMALY
KKSQHTMEVVRRCPEHRCSDGDLAPPHLIRVEGNLYPEYLEDROTFRHSVVVPE
PPEAGSEYTIHYKYMCSNCSMGGMNRPIITITLEDSSGNILGRSPFVAVCACPG
RDRRTTEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLKRGKR
FEMFRLEINALELKDANATEBSGDSRAHSYLTKTKKGQSTSRHKKTVMKKVGPDS"

ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1305;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGGAATCCTCCAT 18
|||||
Db 128 CGACTGACTCTCCAT 111

RESULT 34
MUSP53/c 1322 bp mRNA linear ROD 27-APR-1993
LOCUS Mouse p53 mRNA, complete cds, clone p53-m8.
DEFINITION M13874
M13874.1 GI:200202
p53 gene.
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1322)
AUTHORS Araki, R., Nomura, D., Yokota, K., Wolf, D., Bill, E., Shohat, O. and
Rottier, V.
TITLE Immunologically distinct p53 molecules generated by alternative
splicing
JOURNAL Mol. Cell. Biol. 6 (9), 3232-3239 (1986)
MEDLINE 87064640
PubMed 3023970
REFERENCE

2 (sites)
Hsu, K.A. and Kulesz-Martin, M.F.
TITLE Alternatively spliced p53 RNA in transformed and normal cells of
different tissue types
JOURNAL Nucleic Acids Res. 20 (8), 1979-1981 (1992)
MEDLINE 92253421
PubMed 1579500
COMMENT Original source text: Mouse Ab-MuLV transformed Meth A fibroblast
and normal T-cell line L12, CDNA to mRNA, clone p53-m8.
Location/Qualifiers

FEATURES

source

1..1322
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/map="11"
/cell_line="normal T-cell line L12"
/cell_type="Ab-MuLV transformed Meth A fibroblast"
1..1322
/gene="p53"
/gene="p53"
/codon_start=1
/protein_id="AAA39883.1"
/db_xref="GI:200203"
/translation="MTAMESOSDISLEPLSGEFTSGMLKLLPREDILPSPHCMIDL
LLPDVFEFEGPSEALRVSGAPAAQDPVTETPPVAPAPATPPLSPVPSQKTYG
NYGHLGFLSGTAKSVCTYSPPLNKLFCQIAKTCVQLVWSATPPAGSRVAMALY
KKSQHTMEVVRRCPEHRCSDGDLAPPHLIRVEGNLYPEYLEDROTFRHSVVVPE
PPEAGSEYTIHYKYMCSNCSMGGMNRPIITITLEDSSGNILGRSPFVAVCACPG
RDRRTTEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLKRGKR
FEMFRLEINALELKDANATEBSGDSRAHSYLTKTKKGQSTSRHKKTVMKKVGPDS"

ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1322;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18
 |||||
Db 80 CGACTGGAATCCTCCAT 63

RESULT 35
AB017815/c 1360 bp mRNA linear ROD 09-JUL-1999
LOCUS AB017815
DEFINITION Mus musculus mRNA for p53, complete cds.
ACCESSION AB017815
VERSION AB017815.1 GI:5421807
KEYWORDS p53.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Araki, R., Fukumura, R., Fujimori, A., Tatemura, K. and Abe, M.
TITLE 1 (sites)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1360)
AUTHORS Fujimori, A. and Abe, M.
TITLE Direct Submission
JOURNAL Submitted (22-SEP-1998) Masumi Abe, National Institute of
 Radiological Sciences, Dept. of Biology and Oncology, Anagawa
 4-9-1, Inage-Ku, Chiba, Chiba 263-8555, Japan
 (E-mail: abemasa@nres72.nirs.go.jp, Tel:043-206-3219,
 Fax:043-251-4593)
FEATURES Location/Qualifiers
 source 1..1360
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /cell_line="SR-1"
 /cell_type="epithelial"
 /tissue_type="mammary carcinoma"
 102..1374
 /codon_start=1
 /product="p53"
 /protein_id="BAA82339.1"
 /db_xref="GI:5421808"
 /translation="MTMERSOSDISLEPLSOFPSGLWKLPPEDILPSPHCMDDL
 LIPDVEPPEPGEALRVSGAPADPVTETPGWAPAPATPWPISFVPSOKYOGN
 YGRLGFLGSGTAKSVCTTSPPLNKLFCQLATCPVQLWVSATPPGSRVRAATIK
 KSGQMTVEVRCPEHRCSDGDIAPQHLRVGNLYPEYLRQTFRHSVVVPEP
 PEAGSEYTTIHYKMCNCSGCMGNRRPILITITLEDPSGNLGRDSFEVAVACRG
 DRRTBEENFRKEVLCELPFGSAKRALPTCTASPPQKKPLDGEYFTLKIRGRK
 FEMRELNEALELMDAHTESGDSRAHSYLTKKKGQSTSRHKKTWVKVGPDS"
CDS

ORIGIN
 Query Match 91.1%; Score 16.4; DB 10; Length 1360;
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18
 |||||
Db 128 CGACTGGAATCCTCCAT 111

RESULT 36
MMPS3/c 1377 bp mRNA linear ROD 12-SEP-1993
LOCUS MMPS3
DEFINITION Mouse mRNA for transformation associated protein p53.
ACCESSION X00741
VERSION X00741.1 GI:53570
KEYWORDS oncogene.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 1377)
AUTHORS Jenkins, J.R., Rudge, K., Redmond, S. and Wade-Evans, A.

TITLE Cloning and expression analysis of full length mouse cDNA sequences
 encoding the transformation associated protein p53
JOURNAL Nucleic Acids Res. 12 (14), 5609-5626 (1984)
MEDLINE 84272240
PUBMED 6379601
COMMENT Data kindly reviewed (19-FEB-1986) by A. Wade-Evans.
FEATURES Location/Qualifiers
 source 1..1377
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 123..1292
 /note="unnamed protein product; p53"
 /codon_start=1
 /protein_id="CAA25323.1"
 /db_xref="GI:53571"
 /db_xref="GSA:P02340"
 /db_xref="Swiss-Prot:P02340"
 /translation="MTMERSOSDISLEPLSOFPSGLWKLPPEDILPSPHCMDDL
 LIPDVEPPEPGEALRVSGAPADPVTETPGWAPAPATPWPISFVPSOKYOGN
 YGRLGFLGSGTAKSVCTTSPPLNKLFCQLATCPVQLWVSATPPGSRVRAATIK
 KSGQMTVEVRCPEHRCSDGDIAPQHLRVGNLYPEYLRQTFRHSVVVPEP
 PEAGSEYTTIHYKMCNCSGCMGNRRPILITITLEDPSGNLGRDSFEVAVACRG
 DRRTBEENFRKEVLCELPFGSAKRALPTCTASPPQKKPLDGEYFTLKIRGRK
 FEMRELNEALELMDAHTESGDSRAHSYLTKKKGQSTSRHKKTWVKVGPDS"
ORIGIN
 Query Match 91.1%; Score 16.4; DB 10; Length 1377;
 Best Local Similarity 94.4%; Pred. No. 2.1e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACTGGAATCCTCCAT 18
 |||||
Db 149 CGACTGGAATCCTCCAT 132

RESULT 37
AF151353/c 1409 bp mRNA linear ROD 14-JUL-1999
LOCUS AF151353
DEFINITION Mus musculus tumor suppressor p53 (p53) mRNA, complete cds.
ACCESSION AF151353
VERSION AF151353.1 GI:5081782
KEYWORDS

ORIGIN
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1409)
 Jimenez, G.S., Bryntesson, F., Torres-Arzuayus, M.I., Priestley, A.,
 Beeche, M., Saito, S., Sakaguchi, K., Appella, E., Jeggo, P.A.,
 Taccioli, G.B., Wahl, G.M. and Hubank, M.
 DNA-dependent protein kinase is not required for the p53-dependent
 response to DNA damage
 Nature 400 (6739), 81-83 (1999)
MEDLINE 99330047
PUBMED 10403253
 2 (bases 1 to 1409)
 Jimenez, G.S., Bryntesson, F., Torres-Arzuayus, M.I., Priestley, A.,
 Beeche, M., Saito, S., Sakaguchi, K., Appella, E., Jeggo, P.A.,
 Taccioli, G.B., Wahl, G.M. and Hubank, M.
 Direct Submission
 Submitted (13-MAY-1999) Trafford Centre, University of Sussex,
 Falmer, Brighton, Sussex BN1 9R1, United Kingdom
FEATURES Location/Qualifiers
 source 1..1409
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /map="39.0 cM"
 /cell_line="SGR11"
 /cell_type="fibroblast"
 1..1409
 gene

```

CDS
    /gene="p53"
    /123..1295
    /gene="p53"
    /codon_start=1
    /product="tumor suppressor p53"
    /protein_id="AAD39535.1"
    /db_xref="GI:5081783"
    /translation="MTAMBSQSDISLEPLPSQETPSGLMKLLPPEDILPSPHCMDDL
    LLPDVEEPFEPGSEALRVSGAPAAQDPVTEITPGVAPAPATPMPPLSFVPSQKTYQG
    NYGPHLGFLOSSTAKSVCTYSPPLNKLFCOLAKTCVQLWVATPMPPLSFVPSQKTYQG
    KKSQMTVEVRRCPHHERCSDDGLAPPOHRIYVEGNLYPEYLEDROTFRRSVVAVPE
    PPEAGSEYTTIHYKVMCNSSCMGGMNRRPILITITLEDSSGNLGRDSFEVRVACPG
    RDRRTBEENFRKKEVLCPELPPGSAKRALPTCTASPPQKKPLDGEYFTLKIRGRK
    FEMPRELNEALELKDAAHATESGDSRAHSSYLTKKQGSTSRHKTKVKKVGPDS"
    694
    /gene="p53"
    /replace="c"

variation

ORIGIN
Query Match      91.1%; Score 16.4; DB 10; Length 1409;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY
1 CGACTGTGAATCCTCCAT 18
|||||
149 CGACTGTGACTCTCCAT 132

RESULT 38
AB021961/c 1429 bp mRNA linear ROD 14-APR-2000
LOCUS AB021961
DEFINITION Mus musculus mutant p53 mRNA, complete cds.
ACCESSION AB021961
VERSION AB021961.1 GI:5421849
KEYWORDS p53.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1429)
AUTHORS Araki, R., Fukumura, R., Fujimori, A., Tatsumi, K. and Abe, M.
TITLE Cell cycle in DNA-PKcs knock-out mice
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1429)
AUTHORS Fujimori, A. and Abe, M.
TITLE Direct Submision
JOURNAL Submitted (28-DEC-1998) Masumi Abe, National Institute of
Radiological Sciences, Dept. of Biology and Oncology; Anagawa
4-9-1, Inage-ku, Chiba, Chiba 263-8555, Japan
(E-mail: abemasum@nrxs72.nirs.go.jp, Tel:043-206-3219,
Fax:043-251-4593)
FEATURES
source
    location/Qualifiers
    1..1429
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="SCID"
    /db_xref="taxon:10090"
    /cell_line="SCGR-11"
    /cell_type="fibroblast"
    /note="SCGR-11 cell is a derivative of SC3T3 fibroblast
    cell line derived from scid (severe combined immune
    deficiency) mice."
    1..1429
    /gene="p53"
    /101..1273
    /gene="p53"
    /note="Amino acid no.191 leusine (L) in wild-type mouse
    p53 is substituted to arginine (R)."
```

```

variation
    /gene="p53"
    /note="c in wild-type mouse p53"
    /replace="c"

ORIGIN
Query Match      91.1%; Score 16.4; DB 10; Length 1429;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY
1 CGACTGTGAATCCTCCAT 18
|||||
127 CGACTGTGACTCTCCAT 110

RESULT 39
AB020317/c 1457 bp mRNA linear ROD 09-JUL-1999
LOCUS AB020317
DEFINITION Mus musculus mRNA for p53, complete cds.
ACCESSION AB020317
VERSION AB020317.1 GI:5421827
KEYWORDS p53.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1457)
AUTHORS Fujimori, A. and Abe, M.
TITLE Published Only in Database (1999)
REFERENCE 2 (bases 1 to 1457)
AUTHORS Fujimori, A. and Abe, M.
TITLE Direct Submision
JOURNAL Submitted (19-NOV-1998) Masumi Abe, National Institute of
Radiological Sciences, Dept. of Biology and Oncology; Anagawa
4-9-1, Inage-ku, Chiba, Chiba 263-8555, Japan
(E-mail: abemasum@nrxs72.nirs.go.jp, Tel:043-206-3219,
Fax:043-251-4593)
FEATURES
source
    location/Qualifiers
    1..1457
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="129SvJ"
    /db_xref="taxon:10090"
    /tissue_type="lung fibroblast"
    1..1457
    /gene="p53"
    /102..1274
    /gene="p53"
    /codon_start=1
    /product="p53"
    /protein_id="BAA82343.1"
    /db_xref="GI:5421828"

ORIGIN
Query Match      91.1%; Score 16.4; DB 10; Length 1457;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY
1 CGACTGTGAATCCTCCAT 18
|||||
```


Db 128 CGACTGATCTCTCCAT 111

RESULT 40
AY212017/c

1746 bp mRNA linear ROD 16-MAR-2003

DEFINITION Mus musculus tumor suppressor p53 mRNA, complete cds; alternatively

spliced.

VERSION AY212017.1 GI:28975326

KEYWORDS Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Han, K.A. and Kulesz-Martin, M.F.

TITLE Alternatively spliced p53 RNA in transformed and normal cells of different tissue types

JOURNAL Nucleic Acids Res. 20 (8), 1979-1981 (1992)

MEDLINE 92253421

PUBMED 1579500

REFERENCE 2 (bases 1 to 1746)

AUTHORS Kulesz-Martin, M.F., Lissafeld, B., Huang, H., Kistiel, N.D. and Lee, L.

TITLE Endogenous p53 protein generated from wild-type alternatively

JOURNAL Mol. Cell. Biol. 14 (3), 1698-1708 (1994)

MEDLINE 94158842

PUBMED 8114705

REFERENCE 3 (bases 1 to 1746)

AUTHORS Mu, Y., Liu, Y., Lee, L., Miner, Z. and Kulesz-Martin, M.

TITLE Wild-type alternatively spliced p53: binding to DNA and interaction

JOURNAL EMBO J. 13 (20), 4823-4830 (1994)

MEDLINE 95045373

PUBMED 7957051

REFERENCE 4 (bases 1 to 1746)

AUTHORS Miner, Z. and Kulesz-Martin, M.

TITLE DNA binding specificity of proteins derived from alternatively

JOURNAL Nucleic Acids Res. 25 (7), 1319-1326 (1997)

MEDLINE 97214022

PUBMED 9060424

REFERENCE 5 (bases 1 to 1746)

AUTHORS Kulesz-Martin, M.F., Han, K.A., Lee, L. and Liu, Y.

TITLE Direct Submission

JOURNAL Submitted (02-JAN-2003) Dermatology, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland, OR 97239, USA

FEATURES

SOURCE

1.1746

/organism="Mus musculus"

/mol_type="mRNA"

/strain="BALB/c"

/db_xref="taxon:10090"

/cell_type="non-tumorigenic epidermal cells"

112..1257

/note="alternatively spliced; similar to tumor cell-derived mutant sequence in M13874; p53as"

/codon_start=1

/product="tumor suppressor p53"

/protein_id="AA060156.1"

/db_xref="GI:28975327"

/translation="MTAMESQSDISLPLPSQETPSGLMKLPPEDILPSBPHCMDDL

LLPDVEFFRGPBALRVSGAPAAQDVTETPGVAPATPWPPLSFVPSQKYQG

NYGFLGLOSTAKSVCTSPPLNKLFCQLAKTCPVQWVSATPPAGSVRAAIY

KKSQHTMEVVRCPHERCSDDGLAPQHLIRVGNLYPEYLEDROTFRHSVVVAY

PPBAGSEYTTTHYKMCNCSGCMGNRRPILITLTLEDSGNLGRDSFEYVACPG

RDRRTRENPFRKVCVLCPELPGSAKRALPTCTSAAPPOKKKPLDGGYFTLKIRGRK

FEMRELNALALDKDAHTESGDSRAHSSTLPQRAFOALIRESPNC"

Best Local Similarity 94.4%; Pred. No. 2e+02; Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

1 CGACTGATCTCTCCAT 18

Db 138 CGACTGATCTCTCCAT 121

RESULT 41

MMP53R/c 1773 bp mRNA linear ROD 12-SEP-1993

LOCUS Mouse mRNA for cellular tumour antigen p53.

DEFINITION X01237 K01700

ACCESSION X01237.1 GI:53575

VERSION antigen; tumor antigen.

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Zakut-Houri, R., Oren, M., Biern, B., Lavie, V., Hazum, S. and Givol, D.

TITLE A single gene and a pseudogene for the cellular tumour antigen p53

JOURNAL Nature 306 (5943), 594-597 (1983)

MEDLINE 84068204

PUBMED 6646235

REFERENCE

AUTHORS Biern, B., Zakut-Houri, R., Givol, D. and Oren, M.

TITLE Analysis of the gene coding for the murine cellular tumour antigen

p53

JOURNAL EMBO J. 3 (9), 2179-2183 (1984)

MEDLINE 85027173

PUBMED 6092064

REFERENCE

AUTHORS Givol, R.

TITLE Direct Submission

JOURNAL Submitted (28-NOV-1985)

COMMENT

On Oct 22, 2003 this sequence version replaced gi:200204.

FEATURES

SOURCE

CDS

1.1773

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/note="unnamed protein product; p53 polypeptide (aa

1-390)"

/codon_start=1

/protein_id="CAA25625.1"

/db_xref="GI:53576"

/db_xref="GOA:P02340"

/translation="MTAMESQSDISLPLPSQETPSGLMKLPPEDILPSBPHCMDDL

LLPDVEFFRGPBALRVSGAPAAQDVTETPGVAPATPWPPLSFVPSQKYQG

NYGFLGLOSTAKSVCTSPPLNKLFCQLAKTCPVQWVSATPPAGSVRAAIY

KKSQHTMEVVRCPHERCSDDGLAPQHLIRVGNLYPEYLEDROTFRHSVVVAY

PPBAGSEYTTTHYKMCNCSGCMGNRRPILITLTLEDSGNLGRDSFEYVACPG

RDRRTRENPFRKVCVLCPELPGSAKRALPTCTSAAPPOKKKPLDGGYFTLKIRGRK

FEMRELNALALDKDAHTESGDSRAHSSTLPQRAFOALIRESPNC"

old_sequence

old_sequence

old_sequence

old_sequence

old_sequence

Query Match

91.1%; Score 16.4; DB 10; Length 1746;

/note="G is U in [1]"
 /citation=[1]
 1445. .1446
 /note="U is CU in [1]"
 /citation=[1]
 1447. .1448
 /note="CU is UC in [1]"
 /citation=[1]
 1503
 /note="U is A in [1]"
 /citation=[1]
 1558. .1559
 /note="CU is UC in [1]"
 /citation=[1]

Query Match 91.1%; Score 16.4; DB 10; Length 1773;
 Best Local Similarity 94.4%; Pred. No. 2e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAACTCTCCAT 18
 184 CGACTGTGAACTCTCCAT 167

RESULT 42 BC005448/c 1782 bp mRNA linear ROD 29-JUN-2004
 LOCUS BC005448
 DEFINITION Mus musculus transformation related protein 53, mRNA (CDNA clone
 MGC:6174 IMAGE:3599812), complete cds.
 ACCESSION BC005448
 VERSION BC005448.1 GI:13529427
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1782)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Mak, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheet, T.E., Brownstein, M.J., Usdin, T.B., Toshyuki, S.,
 Carrinci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Wortley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, Y., S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shovenko, Y.,
 Bouffard, G.G., Rodriguez, A.C., Grimwood, J., Schmutz, V., Myers, R.M.,
 Butlerfield, V.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE 12477932
 JOURNAL 2 (bases 1 to 1782)
 PUBMED Strausberg, R.
 REFERENCE Direct Submission
 AUTHORS Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 TITLE Gene Collection (MGC), Cancer Genomics Office, National Cancer
 JOURNAL Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgaps-rc@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil@stanford.edu
 Dickson, M., Schmutz, V., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
 Series: IRAX Plate: 11 Row: 1 Column: 4
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6755880.
 Location/Qualifiers

FEATURES
 SOURCE
 1. .1782
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="FVB/N"
 /clone="MGC:6174 IMAGE:3599812"
 /issue_type="Mammary tumor. Metastatic origin-TGF alpha
 model. 10 month old virgin mouse. Taken by biopsy."
 /clone_1ib="NCI CGAP_Mam1"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1. .1782
 /gene="Trp53"
 /note="synonym: p53"
 /db_xref="LocusID:22059"
 /db_xref="MGI:98834"
 129. .1301
 /gene="Trp53"
 /codon_start=1
 /product="transformation related protein 53"
 /protein_id="AAH05448.1"
 /db_xref="GI:13529428"
 /db_xref="LocusID:22059"
 /db_xref="MGI:98834"
 /translation="MTMRSSQSDISLEPLSGEPTFGMKLLPPEIDILPSHCDDL
 LNFQDVEFPGSGEALRVSGAPAADPVTETPPVAPAPAPWPLSPVQKXYOG
 NYGFHFLQSGTAKSVCTYSPPLKPCQAKTCPLQVWVSAAPPASRVANAIY
 KKSQNTETVTRRCPEHRCSDGGLAPPHLIRVEGNLYPEYLEDRTPRHVVVYE
 PPAGSGEYVTRHFKVNCSSCGMMRRPILITITLEDSSGGLRDSFEVVCAPG
 RDRTEENRKEKEVCPPLPGSAKRALPCTSAAPPKKKPLDGEVETLKIRKRR
 FEFRLNLALEKDAHAIRESGDSRAHSYTKTKKGSTSHKKTWAKVGPDS"

ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 1782;
 Best Local Similarity 94.4%; Pred. No. 2e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAACTCTCCAT 18
 155 CGACTGTGAACTCTCCAT 138

RESULT 43 K02110 2130 bp DNA linear ROD 10-MAY-2004
 LOCUS K02110/c
 DEFINITION Mus musculus tumor antigen p53 (p53) pseudogene, complete cds.
 ACCESSION K02110 X01236
 VERSION K02110.1 GI:200206
 KEYWORDS antigen; p53 gene; processed pseudogene; tumor antigen.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2130)
 AUTHORS Zakut-Houri, R., Oren, M., Bizen, B., Lavie, V., Hazun, S. and Givol, D.
 TITLE A single gene and a pseudogene for the cellular tumour antigen p53
 JOURNAL Nature 306 (5943), 594-597 (1983)
 MEDLINE 84068204
 PUBMED 6646235
 REFERENCE 2 (bases 1 to 2130)

AUTHORS Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Glivol, D.
 TITLE Direct Submission
 JOURNAL Submitted (27-APR-1993) Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot 76100, Israel
 COMMENT On May 10, 2004 this sequence version replaced gi:53574.
 FEATURES
 source
 1..2130
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /clone="pCH53-11"
 <1..1826
 /gene="p53"
 /note="processed pseudogene"
 /pseudo
 1..185
 /gene="p53"
 /note="no similarity to p53 cDNA"
 /pseudo
 169..181
 /rpt_type=direct
 262..1434
 /gene="p53"
 /note="tumor antigen p53"
 /pseudo
 /codon_start=1
 1807..1812
 /gene="p53"
 /note="putative"
 1826
 /gene="p53"
 /note="putative"
 1852..1864
 /rpt_family="direct"

ORIGIN
 Query Match 91.1%; Score 16.4; DB 10; Length 2130;
 Best Local Similarity 94.4%; Pred. No. 2e+02; 1; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 CGACTGATCTCTCCAT 18
 |||||
 Db 288 CGACTGATCTCTCCAT 271

RESULT 44
 AC023805/c
 LOCUS
 DEFINITION Mus musculus clone RP23-402K16, *** SEQUENCING IN PROGRESS ***, 26
 unordered pieces.
 AC023805
 AC023805.7 GI:16118068
 HTG; HTGS PHASE1.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 86146)
 Deckerich, D., Thomas, S., Okumura, G., Carllock, C., Garner, T.,
 Metzger, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,
 Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J.,
 Bunay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,
 Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S.,
 Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R.,
 Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hogues, M.,
 Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,
 Kovari, C., Liu, J., Liu, W., Louissegh, H., Lozano, R., Martin, R.,
 Massey, B., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S.,
 Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogun, M., Parikh, B.,
 Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S.,
 Williamson, A., Wrenford, G., Zhou, X., Bouck, J., Hodgson, A.,
 Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstein, G.,

Worley, K. and Gibbs, R.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 86146)
 AUTHOR WORLEY, K.C.
 TITLE Direct Submission
 JOURNAL Submitted (18-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Oct 14, 2001 this sequence version replaced gi:11079366.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: MABG
 Center clone name: RP23-402K16
 ----- Summary Statistics
 Sequencing vector: M13; L08821
 Chemistry: Dye-primer; Bolyard; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 80653 bases at least Q40
 Consensus quality: 102703 bases at least Q40
 Consensus quality: 112226 bases at least Q20
 Estimated insert size: 102605; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-gel estimation
 Quality coverage: 1.4x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 26 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 4737: contig of 4737 bp in length
 * 4738 4837: gap of unknown length
 * 4838 9473: contig of 4636 bp in length
 * 9474 9573: gap of unknown length
 * 9574 13415: contig of 3842 bp in length
 * 13416 13515: gap of unknown length
 * 13516 17327: contig of 3811 bp in length
 * 17328 17426: gap of unknown length
 * 17427 23207: contig of 5781 bp in length
 * 23208 23307: gap of unknown length
 * 23308 27687: contig of 4380 bp in length
 * 27688 27787: gap of unknown length
 * 27788 31391: contig of 3604 bp in length
 * 31392 31491: gap of unknown length
 * 31492 35716: contig of 4425 bp in length
 * 35717 35816: gap of unknown length
 * 35817 40730: contig of 4914 bp in length
 * 40731 40830: gap of unknown length
 * 40831 43983: contig of 3153 bp in length
 * 43984 44083: gap of unknown length
 * 44084 47104: contig of 3021 bp in length
 * 47105 47205: gap of unknown length
 * 47206 50511: contig of 3307 bp in length
 * 50512 50611: gap of unknown length
 * 50612 53734: contig of 3123 bp in length
 * 53735 53835: gap of unknown length
 * 53836 56324: contig of 2490 bp in length
 * 56325 56424: gap of unknown length
 * 56425 59000: contig of 2576 bp in length
 * 59001 59100: gap of unknown length
 * 59101 61416: contig of 2316 bp in length
 * 61417 61517: gap of unknown length
 * 61518 64430: contig of 2914 bp in length
 * 64431 64530: gap of unknown length

64531 67643: contig of 3113 bp in length
 67744 67744: gap of unknown length
 69936 69936: contig of 2193 bp in length
 70036 70036: gap of unknown length
 72722 72722: contig of 2686 bp in length
 72823 72823: gap of unknown length
 74862 74862: contig of 2040 bp in length
 74963 74963: gap of unknown length
 77123 77123: contig of 2161 bp in length
 77224 77224: gap of unknown length
 79432 79432: contig of 2209 bp in length
 79532 79532: gap of unknown length
 81558 81558: contig of 2026 bp in length
 81659 81659: gap of unknown length
 84009 84009: contig of 2351 bp in length
 84010 84010: gap of unknown length
 86146 86146: contig of 2037 bp in length.

Location/Qualifiers
 1. 86146
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-402K16"

ORIGIN
 Query Match 91.1%; Score 16.4; DB 2; Length 86146;
 Best Local Similarity 94.4%; Pred. No. 1.6e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
 |||||
 Db 10809 CGACTGTGAATCTCCAT 10792

RESULT 45
 LOCUS AL731687 116988 bp DNA linear ROD 06-NOV-2002
 DEFINITION Mouse DNA sequence from clone RP23-56120 on chromosome 11, complete
 sequence.
 ACCESSION AL731687
 VERSION AL731687
 KEYWORDS HMG.
 SOURCE HMG.
 ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 116988)
 Tromans, A.
 Direct Submission
 Submitted (06-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Nov 8, 2002 this sequence version replaced gi:24366489.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-56120 is from the RPCI-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6
 Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

Location/Qualifiers
 1. 11698
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-56120"
 /clone_id="RPCI-23"

ORIGIN
 Query Match 91.1%; Score 16.4; DB 10; Length 116988;
 Best Local Similarity 94.4%; Pred. No. 1.6e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACTGTGAATCTCCAT 18
 |||||
 Db 81044 CGACTGTGAATCTCCAT 81027

RESULT 46
 LOCUS AL445243 150937 bp DNA linear PRI 03-JUL-2001
 DEFINITION Human DNA sequence from clone RP11-538C21 on chromosome 13 contains
 STSs and GSSs, complete sequence.
 ACCESSION AL445243
 VERSION AL445243
 KEYWORDS HMG.
 SOURCE Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 150937)
 Wall, M.
 Direct Submission
 Submitted (29-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Nov 10, 2000 this sequence version replaced gi:11024525.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone configs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13>

IMPORTANT: This sequence is not the entire insert of clone RP11-538C21. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-538C21 is at 1 in this sequence. The true left end of clone RP11-168U5 is at 150838 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an

alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-538C21 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.

FEATURES

```
source
1..150937
/organism="Homo sapiens"
/db_type="genomicDNA"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-538C21"
/clone_id="RP11-11.2"
896..951
repeat_region
/note="28 copies 2 mer tt 69% conserved"
3412..3905
misc_feature
/note="match: GSS: Em:A0663918"
3949..4172
repeat_region
/note="112 copies 2 mer tt 67% conserved"
complement(5225..5632)
/note="match: GSS: Em:A0012043"
complement(5290..5643)
/note="match: GSS: Em:A0056175"
complement(6845..7233)
/note="match: GSS: Em:AQ211228"
7242..7762
misc_feature
/note="match: GSS: Em:A2519265"
complement(8525..9078)
/note="match: GSS: Em:AQ422480"
complement(8631..9064)
/note="match: STS: Em:G53065
match: GSS: Em:B76751"
10605..10640
repeat_region
/note="18 copies 2 mer aa 88% conserved"
19650..19717
repeat_region
/note="34 copies 2 mer tc 85% conserved"
complement(28596..28945)
/note="match: GSS: Em:AQ458456"
29096..29557
misc_feature
/note="match: GSS: Em:AQ772130"
complement(30788..31290)
/note="match: STS: Em:G57193
match: GSS: Em:AQ341794"
complement(34113..34607)
/note="match: GSS: Em:AQ194187"
complement(34157..34604)
/note="match: GSS: Em:AQ814840"
complement(34168..34607)
/note="match: GSS: Em:AQ198600"
34713..34802
repeat_region
/note="WT2FA repeat: matches 81..181 of consensus"
38474..38849
repeat_region
/note="THB1C repeat: matches 1..371 of consensus"
41672..42240
repeat_region
/note="THB1B-INTERNAL repeat: matches 1..615 of consensus"
44388..44427
repeat_region
/note="20 copies 2 mer ac 95% conserved"
complement(44392..44726)
/note="match: GSS: Em:AQ488076"
44846..45259
misc_feature
/note="match: GSS: Em:B90508"
complement(45859..46009)
/note="match: STS: Em:AP003550"
46158..46418
repeat_region
/note="MERS2A repeat: matches 528..750 of consensus"
46440..46617
repeat_region
/note="LTR28 repeat: matches 1..178 of consensus"
47959..48082
misc_feature
/note="match: STS: Em:G27256"
```

```
repeat_region
48084..48147
/note="32 copies 2 mer ta 70% conserved"
48088..48301
misc_feature
/note="match: STS: Em:G27256"
49323..49712
misc_feature
/note="match: GSS: Em:AQ375299"
53872..53931
repeat_region
/note="MER67D repeat: matches 445..506 of consensus"
54086..54442
repeat_region
/note="MER67C repeat: matches 1..367 of consensus"
54824..55503
misc_feature
/note="match: GSS: Em:AQ055122"
57402..57435
repeat_region
/note="17 copies 2 mer ct 91% conserved"
63756..64537
misc_feature
/note="match: GSS: Em:AQ537963"
complement(63874..64285)
/note="match: GSS: Em:AQ465707"
66074..66115
repeat_region
/note="21 copies 2 mer tt 85% conserved"
67279..67780
misc_feature
/note="match: GSS: Em:AQ765920"
complement(68370..68622)
/note="match: GSS: Em:A2391162"
complement(68391..69107)
/note="match: GSS: Em:AQ527862"
complement(68437..68626)
/note="match: GSS: Em:AQ807981"
complement(68514..68639)
/note="match: GSS: Em:AQ456087"
69124..69559
misc_feature
/note="match: GSS: Em:AQ881502"
71416..71505
repeat_region
/note="45 copies 2 mer ta 73% conserved"
71511..71871
repeat_region
/note="L1P3 repeat: matches 5783..6146 of consensus"
78843..79653
repeat_region
/note="T1ger3b repeat: matches 407..1231 of consensus"
79664..80405
repeat_region
/note="T1ger3b repeat: matches 1..407 of consensus"
82762..83871
repeat_region
/note="L1P15 repeat: matches 604..1684 of consensus"
83872..84741
repeat_region
/note="MER1D repeat: matches 3..897 of consensus"
84742..86381
repeat_region
/note="L1P15 repeat: matches 1684..3280 of consensus"
86682..89269
repeat_region
/note="L1P15 repeat: matches 3280..6157 of consensus"
91758..91799
repeat_region
/note="21 copies 2 mer ac 95% conserved"
91815..91870
repeat_region
/note="28 copies 2 mer ac 78% conserved"
92029..92491
repeat_region
/note="LTR33 repeat: matches 38..521 of consensus"
95535..96172
repeat_region
/note="MER76 repeat: matches 1..666 of consensus"
96190..96217
repeat_region
/note="14 copies 2 mer cc 89% conserved"
96233..96254
repeat_region
/note="MER76 repeat: matches 667..688 of consensus"
97703..98137
misc_feature
/note="match: GSS: Em:AQ077541"
98891..99319
repeat_region
/note="L1P repeat: matches 4766..5195 of consensus"
102030..102067
repeat_region
/note="19 copies 2 mer tt 78% conserved"
109590..109767
repeat_region
/note="89 copies 2 mer tt 66% conserved"
109797..110092
repeat_region
/note="MSTC repeat: matches 1..289 of consensus"
11317..113786
misc_feature
/note="match: GSS: Em:AQ017010"
complement(114392..114852)
```

```

misc_feature      /note="match: GSS: Em:B94847"
                  /note="complement(114424, 116842)"
misc_feature      /note="match: GSS: Em:AQ119585"
                  /note="complement(116203, 116626)"
misc_feature      /note="match: GSS: Em:AQ081553"
                  /note="complement(117291, 117478)"
misc_feature      /note="match: GSS: Em:AQ33498 Em:AQ431610"
                  /note="match: GSS: Em:AZ497064"
misc_feature      /note="match: GSS: Em:AQ319097"
                  /note="complement(117313, 117470)"
misc_feature      /note="match: GSS: Em:G10072"
                  /note="match: GSS: Em:G10072"
repeat_region     /note="54 copies 2 mer ta 64 conserved"

Query Match      91.1%; Score 16.4; DB 9; Length 150937;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CGACTGTGAATCCTCCAT 18
Db 41556 CGACTGTGAATCCTCCAT 41573

```

```

RESULT 47
AC102741 163447 bp DNA linear HTG 20-MAR-2004
AC102741/c Mus musculus chromosome 17 clone RP24-343110 map 17, *** SEQUENCING
DEFINITION
AC102741 GI:45597679
AC102741.4 HTG: HTGS_PHASE1; HTGS_FULFILL; HTGS_ACTIVEFIN.
VERSION
HTG: HTGS_PHASE1; HTGS_FULFILL; HTGS_ACTIVEFIN.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 163447)
Britten, B., Nussbaum, C. and Lander, E.
Mus musculus chromosome 17, clone RP24-343110
Unpublished
2 (bases 1 to 163447)
Britten, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, M., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,
Lamaze, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
Menus, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, W., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Stratus, N., Subramanian, A., Talamas, J., Testfay, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 163447)
Britten, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepe, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearrellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,

```

```

TITLE
JOURNAL
COMMENT

```

```

Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, X., Liu, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
McCarthy, M., Menus, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Stubbbs, M., Talamas, J., Testfay, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (20-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 20, 2004 this sequence version replaced g1:28630066.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center Project name: L19396
Center Clone name: 343_I10

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 16709: contig of 16709 bp in length
* 16710 16809: gap of 100 bp
* 16810 20326: contig of 3517 bp in length
* 20327 20426: gap of 100 bp
* 20427 26699: contig of 6273 bp in length
* 26700 26799: gap of 100 bp
* 26800 51907: contig of 25108 bp in length
* 51908 52007: gap of 100 bp
* 52008 97610: contig of 45603 bp in length
* 97611 97710: gap of 100 bp
* 97711 134460: contig of 36750 bp in length
* 134461 134560: gap of 100 bp
* 134561 150090: contig of 15530 bp in length
* 150091 150190: gap of 100 bp
* 150191 163447: contig of 13257 bp in length.
Location/Qualifiers
1. 163447
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/map="17"
/clone="RP24-343110"
/clone_lib="RPCI-24 Male Mouse BAC"

```

ORIGIN

```

Query Match      91.1%; Score 16.4; DB 2; Length 163447;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGAATCCTCCAT 18
Db 106744 CGACTGTGAATCCTCCAT 106727

```

RESULT 48
BX324147 172568 bp DNA linear HTG 24-MAY-2003
LOCUS
DEFINITION Dario rerio clone DKEY-217G15, WORKING DRAFT SEQUENCE.
ACCESSION BX324147
VERSION BX324147.2 GI:31075194
KEYWORDS HTG, HTGS_PHASE2, HTGS_DRAFT, HTGS_FULFILL.
SOURCE Dario rerio (zebrafish)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
Cypriniformes; Cyprinidae; Dario.
1 (bases 1 to 172568)
McLaren, S.
Direct Submission
Submitted (23-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 25, 2003 this sequence version replaced gi:30024560.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK217G15
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 172418 bases at least Q40
Consensus quality: 172539 bases at least Q30
Consensus quality: 172564 bases at least Q20
Insert size: 172568; sum-of-contigs
Insert size: 162997; 3.0% error; agarose-fp
Quality coverage: 5.65x in Q20 bases; sum-of-contigs Quality
coverage: 6.03x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 172568: contig of 172568 bp in length.
Location/Qualifiers
1..172568
/organism="Dario rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-217G15"
/clone_lib="Daniokey"
1..172568
/note="assembly_fragment:01051"

ORIGIN
Query Match 91.1%; Score 16.4; DB 2; Length 172568;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGACTGTGAATCTTCAT 18
DB 106266 CGACTGTGAATCTTCGT 106283
RESULT 49
AC074146 201602 bp DNA linear HTG 15-JUL-2000
LOCUS
DEFINITION Mus musculus clone RP23-107G3, WORKING DRAFT SEQUENCE, 38 unordered
pieces.

ACCESSION AC074146
VERSION AC074146.1 GI:9211256
KEYWORDS HTG, HTGS_PHASE1, HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 201602)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 201602)
DOE Joint Genome Institute.
Direct Submission
Submitted (15-JUL-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>
----- Project Information
Center Project Name: 0
Center clone name: RP23-107G3

Summary Statistics
Consensus quality: 134464 bases at least Q40
Consensus quality: 159739 bases at least Q30
Consensus quality: 170106 bases at least Q20
Estimated insert size: 181300; agarose-fp estimation
Estimated insert size: 197902; sum-of-contigs estimation
Quality coverage: 3.41 in Q20 bases; agarose-fp estimation
Quality coverage: 3.12 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 38 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1675: contig of 1675 bp in length
1676 1775: gap of unknown length
1776 3291: contig of 1516 bp in length
3292 3391: gap of unknown length
3392 4432: contig of 1041 bp in length
4433 4532: gap of unknown length
4533 6053: contig of 1521 bp in length
6054 6153: gap of unknown length
6154 7161: contig of 1008 bp in length
7162 7261: gap of unknown length
7262 8417: contig of 1156 bp in length
8418 8517: gap of unknown length
8518 9800: contig of 1283 bp in length
9801 9900: gap of unknown length
9901 10926: contig of 1026 bp in length
10927 11026: gap of unknown length
11027 12062: contig of 1036 bp in length
12063 12162: gap of unknown length
12163 14110: contig of 1948 bp in length
14111 14210: gap of unknown length
14211 14211: gap of unknown length
14212 15534: contig of 1324 bp in length
15535 15634: gap of unknown length
15635 17316: contig of 1682 bp in length
17317 17416: gap of unknown length
17417 20982: contig of 3566 bp in length
20983 21082: gap of unknown length
21083 24936: contig of 3854 bp in length
24937 25036: gap of unknown length
25037 28124: contig of 3088 bp in length
28125 28224: gap of unknown length
28225 30948: contig of 2724 bp in length
30949 31048: gap of unknown length

```

* 31049 34011: contig of 2963 bp in length
* 34012 34111: gap of unknown length
* 34112 37382: contig of 3271 bp in length
* 37383 37482: gap of unknown length
* 37483 41136: contig of 3654 bp in length
* 41137 41236: gap of unknown length
* 41237 45754: contig of 4518 bp in length
* 45755 45854: gap of unknown length
* 45855 48474: contig of 2620 bp in length
* 48475 48574: gap of unknown length
* 48575 52059: contig of 3485 bp in length
* 52060 52159: gap of unknown length
* 52160 55993: contig of 3834 bp in length
* 55994 56093: gap of unknown length
* 56094 59136: contig of 3043 bp in length
* 59137 59236: gap of unknown length
* 59237 63015: contig of 3779 bp in length
* 63016 63115: gap of unknown length
* 63116 66917: contig of 3802 bp in length
* 66918 67017: gap of unknown length
* 67018 72225: contig of 5208 bp in length
* 72226 72325: gap of unknown length
* 72326 80793: contig of 8468 bp in length
* 80794 80893: gap of unknown length
* 80894 88129: contig of 7236 bp in length
* 88130 88229: gap of unknown length
* 88230 95766: contig of 7537 bp in length
* 95767 95866: gap of unknown length
* 95867 104692: contig of 8826 bp in length
* 104693 104792: gap of unknown length
* 104793 113453: contig of 8661 bp in length
* 113454 113553: gap of unknown length
* 113554 121609: contig of 8056 bp in length
* 121610 121709: gap of unknown length
* 121710 132117: contig of 10408 bp in length
* 132118 132217: gap of unknown length
* 132218 143836: contig of 11619 bp in length
* 143837 143936: gap of unknown length
* 143937 153823: contig of 9887 bp in length
* 153824 153923: gap of unknown length
* 153924 174899: contig of 20976 bp in length
* 174900 174999: gap of unknown length
* 175000 201602: contig of 26603 bp in length.

FEATURES
  source
    1..201602
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /clone="RP23-107G3"
    /clone_lib="RP23 mouse BAC library 23"

ORIGIN
  Query Match      91.1%; Score 16.4; DB 2; Length 201602;
  Best Local Similarity 94.4%; Pred. No. 1.5e+02;
  Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGACTGATGATCTCCAT 18
    |||||
Db 197993 CGACTGATGATCTCCAT 197976

RESULT 50
AC148732/c AC148732 214262 bp DNA linear HTG 31-MAR-2004
LOCUS AC148732
DEFINITION Callithrix jacchus clone CH259-164P15, WORKING DRAFT SEQUENCE, 11
unordered pieces.
AC148732
AC148732.1 GI:45860953
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Callithrix jacchus (white-tufted-ear marmoset)
SOURCE Callithrix jacchus
ORGANISM Callithrix jacchus
          Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euleleostomi;
          Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;

```

```

REFERENCE
  AUTHORS
    Callithrix.
    1 (bases 1 to 214262)
  Antonelli,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
  Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
  Coleman,H., Daki,N., Engle,J., Grantie,S., Gun,X., Gupta,D.,
  Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
  Idol,J.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Latic,P.,
  Larson,S., Lee-Jin,S.-O., Legaspi,R., Maduro,O.L., Maduro,V.B.,
  Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J.,
  Mullikin,J.C., Paguirigan,C., Portney,M.E., Prasad,A., Puri,O.,
  Redix-Duque,N., Schandler,K., Schueler,M.G., Shah,K., Sison,C.,
  Stantipop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.D.,
  Wehenby,K.D., Young,A. and Green,E.D.
  NISC Comparative Sequencing Initiative
  Unpublished
  2 (bases 1 to 214262)
  Green,E.D.
  Direct Submission
  Submitted (31-MAR-2004) NIH Intramural Sequencing Center, 8717
  Groveomont Circle, Galthersburg, MD 20877, USA
  ----- Genome Center
  Center: NIH Intramural Sequencing Center
  Center code: NISC
  Web site: http://www.nisc.nih.gov
  Contact: nisc_zoemngri.nih.gov
  ----- Project Information
  Center project name: 164P15
  Center clone name: 164P15
  ----- Summary Statistics
  Sequencing vector: plasmid; n/a; 100% of reads
  Chemistry: Dye-terminator Big Dye; 100% of reads
  Assembly program: Phrap; version 0.990319
  Consensus quality: 209757 bases at least Q40
  Consensus quality: 210672 bases at least Q30
  Consensus quality: 211371 bases at least Q20
  Insert size: 213000; agarose-fp
  Insert coverage: 213262; sum-of-contigs
  Quality coverage: 8.19x in Q20 bases; sum-of-contigs
  Quality coverage: 8.18x in Q20 bases; sum-of-contigs
  -----
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 11 contigs. The true order in this sequence record is
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  1
  2648: contig of 2648 bp in length
  2649 2748: gap of unknown length
  2749 4920: contig of 2172 bp in length
  4921 5020: gap of unknown length
  5021 7682: contig of 2662 bp in length
  7683 7782: gap of unknown length
  7783 22178: contig of 14396 bp in length
  22179 22278: gap of unknown length
  22279 29583: contig of 7305 bp in length
  29584 29683: gap of unknown length
  29684 40166: contig of 10483 bp in length
  40167 40266: gap of unknown length
  40267 57686: contig of 17420 bp in length
  57687 57786: gap of unknown length
  57787 73989: contig of 16203 bp in length
  73990 74089: gap of unknown length
  74090 97609: contig of 23520 bp in length
  97610 97709: gap of unknown length
  97710 144022: contig of 46313 bp in length
  144023 144122: gap of unknown length
  144123 214262: contig of 70140 bp in length.
  -----
  FEATURES
    source
      1..214262
      /organism="Callithrix jacchus"
      /mol_type="genomic DNA"

```



```

/db_xref="taxon:9483"
/clone="CH259-164P15"
/clone_lib="CH259"
/note="BAC resource: http://bacpac.chori.org/"
misc_feature
1..2648
/note="assembly_fragment"
misc_feature
2749..4920
/note="assembly_fragment"
misc_feature
5021..7682
/note="assembly_fragment"
misc_feature
7783..22178
/note="assembly_fragment"
misc_feature
22279..29583
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
misc_feature
29684..40166
/note="assembly_fragment"
misc_feature
40267..57686
/note="assembly_fragment"
misc_feature
57787..73989
/note="assembly_fragment"
misc_feature
74090..97609
/note="assembly_fragment"
misc_feature
97710..144022
/note="assembly_fragment"
clone_end:r7
vector_side:left"
misc_feature
144123..214262
/note="assembly_fragment"

```

ORIGIN

```

Query Match          91.1%; Score 16.4; DB 2; Length 214262;
Best Local Similarity 94.4%; Pred.No.1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CGACTGTGATCTCTCCAT 18
        |||||
Db      2777 CGACTGTGATCTCTCCAT 2760

```

Search completed: January 14, 2005, 16:03:08
 Job time : 984.263 secs

THIS PAGE BLANK (USPTO)